

Tesis doctoral

**MODELIZACIÓN ARCH.  
ESTIMACIÓN DE LA  
VOLATILIDAD DEL IBEX-35**

Autor: D. Rafael de Arce Borda  
Director: Dr. Julián Pérez García

Octubre, 2000

T/351

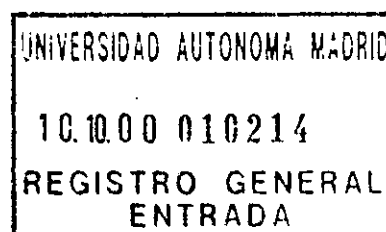


Tesis doctoral

**MODELIZACIÓN ARCH.  
ESTIMACIÓN DE LA  
VOLATILIDAD DEL IBEX-35**

R.º FEE. 77062  
M

Autor: D. Rafael de Arce Borda  
Director: Dr. Julián Pérez García



Octubre, 2000

a 381684

*"Cuando una persona desea realmente algo, el universo entero conspira para que pueda realizar su sueño. Basta con aprender a escuchar los dictados del corazón y a descifrar el lenguaje que está más allá de la palabras, el que muestra aquello que los ojos no pueden ver."*

*El Alquimista, Paulo Coelho (1998)*

*Gracias a ese universo de personas que con su apoyo diario han permitido que concluyera esta tesis, en especial,*

*a mis padres y hermanos, por su apoyo incondicional y por aguantar mis malos humores cuando el cansancio me vence;*

*a mis compañeros de trabajo, por obligarme a aprender día a día y prestarme su ayuda siempre que la necesito;*

*a mis amigos, por estar siempre ahí.*

1. INTRODUCCIÓN	5
1.1 Fundamentos básicos para el desarrollo de los modelos ARCH	9
1.2 Interpretación básica de los modelos de heterocedasticidad condicional	13
1.3. Recorrido histórico y principales variantes de los modelos ARCH	15
2. MODELOS AUTORREGRESIVOS HETEROCEDÁSTICOS CONDICIONALES: ARCH(q)	21
2.1. Especificación de un modelo ARCH(q)	23
2.2. Condición de estacionariedad de un ARCH(q)	24
2.3. Estimación de un ARCH(q)	27
2.4. Modelo ARCH de Regresión	32
2.5. Contrastes de estructura ARCH	36
2.6. Predicción con modelos ARCH	37
2.7. Implicaciones de un arch(1) sobre la estructura de un modelo	41
3. MODELO ARCH EN MEDIA: ARCH-M	45
3.1. Especificación	47
3.2. Estimación	48
4. MODELO GENERALIZADO DE HETEROCEDASTICIDAD CONDICIONAL AUTORREGRESIVA: GARCH(p,q)	51
4. 1. Especificación	52
4.2. Condición de estacionariedad de un proceso GARCH(p,q)	54
4.3. Estimación del Modelo GARCH de regresión	58
4.4. Contraste de estructura GARCH(p,q)	60
4.5. Predicción	64
5. VARIANTES SOBRE LOS MODELOS AUTORREGRESIVOS HETEROCEDÁSTICOS CONDICIONALES	67
5.1. Modelo IGARCH y Modelos de Memoria Larga (Component ARCH)	68
Introducción	68
Modelos de memoria larga o ARCH de componentes	70
5.2. Modelo E-GARCH	78
Introducción	78
Especificación del modelo GARCH exponencial (EGARCH)	79
Contraste de especificación	82
5.3. Modelos GJR - ARCH	84
Introducción	84
Método de estimación por variables instrumentales de Campbell (1987)	87
Modelo GJR-ARCH: Modificaciones sobre el modelo GARCH-M de Bollerslev	88
Estimación y contraste	91
5.4. Modelos de cadenas de Markov y SW-ARCH (Switching-ARCH)	92
Cambios de régimen con cadenas de markov	92
Introducción a las cadenas de Markov	92
Relación de vectores autorregresivos en las cadenas de Markov	94



Modelo IV alternativo: GJR-ARCH + ARCH de memoria larga (Modelo IV + ARCH de componentes) _____	208
Modelo V: GJR-ARCH con estructura E-GARCH (Modelo IV con errores normalizados y en logaritmos) _____	211
Comparación conjunta de los modelos I al V _____	214
Modelo VI: especificación con potencias asimétricas (AP-ARCH) _____	221
Modelo VII: Una propuesta alternativa: modelo GJR introduciendo el óptimo de desfase de correlación tipo APARCH _____	224
Diferencias GJR - AP - ARCH y GJR-ARCH _____	227
Comparación de los modelos presentados y contraste de permanencia estructural del modelo elegido. Análisis de la volatilidad estimada _____	229
9. CONSIDERACIONES FINALES _____	235
ANEXO I: Valores estimados de la varianza condicional autorregresiva y calificación _____	253
ANEXO II: Programas E-views creados para las estimaciones utilizadas en el estudio _____	265
MODELOS TIPO ARCH _____	265
ALGUNAS MEDIDAS DE CONTRASTACIÓN EMPLEADAS _____	276
Referencias Bibliográficas _____	279

# 1.INTRODUCCIÓN

En muchas ocasiones en economía se habla de sucesos condicionados o de generación de expectativas a partir de los movimientos relativos que se produjeron en el pasado. Por ejemplo, todo el mundo relaciona inmediatamente la estabilidad o la inestabilidad en los mercados financieros con su comportamiento inmediatamente anterior, produciéndose fuertes hondas en la evolución de sus variables que, después de un gran sobresalto que dura más o menos días, tienden a retomar una senda de evolución tranquila. A cualquiera se le ocurre entonces que, en variables como éstas, el comportamiento en el momento actual responde a una expectativa generada sobre el valor de cambio producido en el momento precedente; es decir, a un valor esperado condicionado por la varianza del período anterior.

En la teoría clásica de series temporales (metodología de Box-Jenkins), el desarrollo estadístico se realiza a partir de un proceso estocástico estacionario; es decir (en sentido amplio o débil) de un proceso con:

- Media constante.
- Varianza constante.
- correlación entre dos observaciones distintas igual a la de otras dos cualquiera separadas por la misma distancia (mismo número de periodos).

En torno a la confirmación de la ausencia de tendencia (determinista o aleatoria), hay un nutrido conjunto de teorías y desarrollos matemáticos centrados en la diferenciabilidad de la serie temporal y en la existencia o no de raíces unitarias a partir de los conocidos contrastes de Dickey y Fuller, de Mackinon o de Phillips y Perron, por citar algunos. Sin embargo, el estudio de la componente de varianza constante es un fenómeno menos extendido y, no tener en cuenta una posible no constancia de este componente, puede suponer diversos problemas estadísticos cuando se estiman modelos econométricos (problemas ligados con la eficiencia de los parámetros estimados y su fuerte volatilidad ante el amplio intervalo de confianza en el que se mueven).

Determinar un patrón de comportamiento estadístico para la varianza es el cometido de los modelos Autorregresivos condicionales heterocedásticos: ARCH. Engle (1982) es el autor de una primera aproximación a la varianza condicional del tipo que describiremos más adelante. Después de estos hay una amplia familia de sofisticaciones del modelo inicial que darán nombre a los modelos GARCH, IGARCH, E-GARCH, TARCH, SWARCH, QS-ARCH, APARCH, FACTOR-ARCH, ...

En el artículo seminal de los modelos ARCH, Engle cita tres situaciones que motivan y justifican la modelización de la Heterocedasticidad Condicional Autorregresiva (nombre por él mismo dado). Estas serían las siguientes:

1. La experiencia empírica nos lleva a contrastar períodos de amplia varianza de error seguidos de otros de varianza más pequeña. Es decir, el valor de la dispersión del error respecto a su media cambia en el pasado, por lo que es lógico pensar que un modelo que atienda en la predicción a los valores de dicha varianza en el pasado servirá para realizar estimaciones más precisas.
2. En segundo lugar, Engle expone la validez de estos modelos para determinar los criterios de mantenimiento o venta de activos financieros. Los agentes económicos deciden esta cuestión en función de la información proveniente del pasado respecto al valor medio de su rentabilidad y la volatilidad que ésta ha tenido. Con los modelos ARCH se tendrían en cuenta estos dos condicionantes.
3. El modelo de regresión ARCH puede ser una aproximación a un sistema más complejo en el que no hubiera factores innovacionales con heterocedasticidad condicional. Los modelos estructurales admiten, en multitud de ocasiones, una especificación tipo ARCH infinita que determina con parámetros cambiantes, lo que hace a este tipo de modelos capaces de contrastar la hipótesis de permanencia estructural que supone una de las hipótesis de partida y condición necesaria para la validez del modelo econométrico tradicional.

En definitiva, la clave de estos modelos está en considerar la información pasada de la variable y su volatilidad observada como factor altamente explicativo de su comportamiento presente y, por extensión lógica, de su futuro predecible. Estadísticamente, esta conclusión se refleja en tener en cuenta la esperanza condicional (conocida y fija la información hasta el momento inmediatamente anterior) del cuadrado de una variable (la expresión de su varianza si su media es nula).

## 1.1 Fundamentos básicos para el desarrollo de los modelos ARCH

En este apartado, se pretende sintetizar los supuestos teórico-estadísticos básicos que centran la realización de modelos ARCH y sus variantes. Con ello, se quiere no sólo establecer el punto de partida de este tipo de modelos, sino también plantear las ecuaciones sobre las que posteriormente se irán aclarando distintos aspectos de la metodología propuesta por Engle.

Partiendo de los principios del análisis de las series temporales, comenzamos con la definición de un proceso estocástico estacionario como aquella sucesión ordenada de variables aleatorias cuya función de distribución es invariante ante valores igualmente separados:

$$Y_t^\infty | F(y_{t-\infty}, y_{t-1-\infty}, \dots, y_t, y_{t+1}, \dots, y_{t+\infty}) = F(y_{t+m-\infty}, y_{t+m-1-\infty}, \dots, y_{t+m}, y_{t+m+1}, \dots, y_{t+m+\infty})$$

Como en la práctica es casi imposible conocer la verdadera función de distribución de muchos procesos aleatorios, esta definición (que se conoce con el nombre de "estacionariedad en sentido fuerte") se suele confirmar sólo para el primer y los segundos momentos; es decir, para la media y la varianza del proceso. Según esta definición de "estacionariedad en sentido amplio o débil", un proceso estocástico sería estacionario cuando se cumplieran las tres condiciones siguientes:

1.  $E(Y_t) = \mu$ , ó media constante,
2.  $\text{Var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2$ , o varianza constante, y
3.  $\text{Cov}(Y_t, Y_{t-j}) = \text{Cov}(Y_{t+m}, Y_{t+m-j})$ , o covarianza igual para pares de observaciones igualmente distanciados.

Como es conocido, el "ruido blanco" es un caso particular de este tipo de proceso en el que las tres condiciones se reescribirían del siguiente modo:

1.  $E(\varepsilon_t) = 0$
2.  $\text{Var}(\varepsilon_t) = E(\varepsilon_t - 0)^2 = \sigma_\varepsilon^2$
3.  $\text{Cov}(\varepsilon_t, \varepsilon_{t-j}) = 0 \quad \forall j \neq 0$

Sobre esta definición clásica de estacionariedad, conviene hacer algunas puntualizaciones estadísticas relativas a las segundas derivadas del proceso que estamos manejando:

A. El hecho de que no exista autocorrelación entre observaciones del ruido blanco desplazadas en el tiempo, no significa necesariamente que no haya dependencia entre estas de un modo no lineal. Lo único que aseguramos es precisamente eso: no podemos formular ningún tipo de dependencia lineal entre  $\varepsilon_t$  y  $\varepsilon_{t-j}$ , pero nada se dice al respecto de si puede haber una relación de dependencia cuadrática, exponencial o de cualquier otro tipo.

B. La lógica de la dependencia entre el ritmo de evolución en periodos precedentes y el valor de variación del periodo actual nos lleva necesariamente a hablar de probabilidades condicionadas en términos de estadística teórica o inferencial. Es a partir de los momentos de primer y segundo orden en términos condicionales como se pueden descubrir relaciones de causalidad entre series temporales que responden como un proceso estocástico estacionario para el contraste lineal.

En definitiva, puede existir un proceso definido a partir de un "ruido blanco", en el que la media y la varianza marginales sean constantes; y, al mismo tiempo, la media condicional puede ser constante y la varianza condicional no fija.

En el siguiente proceso:

$$y_t = \varepsilon_t (\alpha_0 + \alpha_1 y_{t-1}^2)^{1/2} \quad \text{Ec. 1}$$

- donde  $\varepsilon_t$  es un proceso de "ruido blanco" (entre otras, no hay correlación con su pasado, luego tampoco la hay con el pasado de  $y_t$ ).
- El proceso generado  $y_t$  es también estacionario
- En los momentos condicionales, en "t", el valor de "t-1" es una realización concreta conocida (no aleatoria)

Podemos calcular los valores de su esperanza y su varianza, no olvidando que, en los valores condicionales, los valores de la endógena en t-1 son realizaciones ya

conocidas del proceso aleatorio, por lo que su valor esperando es el valor ya observado (no son variable aleatoria):

Cuadro 1: Momentos marginales y condicionales

	<b>Marginal (incondicional)</b>	<b>Condicional</b>
<b>Esperanza</b>	$E(y_t) = E(\varepsilon_t^2 (\alpha_0 + \alpha_1 y_{t-1}^2))^{1/2} = E(\varepsilon_t) E((\alpha_0 + \alpha_1 y_{t-1}^2)^{1/2}) = 0$	$E_{t-1}(y_t) = (\alpha_0 + \alpha_1 y_{t-1}^2)^{1/2} E_{t-1}(\varepsilon_t) = 0$
<b>Varianza</b>	$E(y_t^2) = E(\varepsilon_t^2 (\alpha_0 + \alpha_1 y_{t-1}^2)) = \sigma_\varepsilon^2 (\alpha_0 + \alpha_1 E(y_{t-1}^2));$ $h_t = E(y_t^2) = \frac{\alpha_0}{1 - \alpha_1} h_\varepsilon$	$E_{t-1}(y_t^2) = (\alpha_0 + \alpha_1 y_{t-1}^2) E_{t-1}(\varepsilon_t^2);$ $\sigma_t^2 = (\alpha_0 + \alpha_1 y_{t-1}^2) h_\varepsilon$

Como resumen de lo expuesto en el Cuadro 1, es interesante resaltar que:

- la media (la esperanza) es contante en ambos casos, e igual a cero.
- La varianza marginal es constante; mientras que
- la varianza condicional depende de los valores que haya tomado  $y_{t-1}^2$ , luego no es fija.

Se da la circunstancia de que, en este proceso, la función de autocovarianza es nula para todos los retardos que se quieran considerar; mientras que es distinta de cero para los valores al cuadrado del mismo proceso  $y_t$  generado.

La autocovarianza de  $y_t$  para el retardo  $\tau$  se podría calcular como:

$$\gamma(\tau) = E(y_t y_{t-\tau}) = E(\varepsilon_t (\alpha_0 + \alpha_1 y_{t-1}^2)^{1/2} y_{t-\tau}) = E(\varepsilon_t) E((\alpha_0 + \alpha_1 y_{t-1}^2)^{1/2} y_{t-\tau}) = 0 \quad \text{Ec. 2}$$

dado que el proceso lo hemos definido como estacionario y el residuo como "ruido blanco"; es decir, con correlación nula con periodos precedentes, lo que implica también correlación nula con los valores de  $y_t$  desplazados en el tiempo.

La función de autocovarianza para la serie al cuadrado presenta valores distintos de cero, por lo menos para el retardo uno, como se puede comprobar con la siguiente expresión:

$$\begin{aligned} \text{cov}(y_t^2, y_{t-1}^2) &= E((y_t^2 - E(y_t^2))(y_{t-1}^2 - E(y_{t-1}^2))) = E(\varepsilon_t^2) E\left((\alpha_0 + \alpha_1 y_{t-1}^2) - \frac{\alpha_0}{1-\alpha_1}\right) \left(y_{t-1}^2 - \frac{\alpha_0}{1-\alpha_1}\right) \\ &= \left(\alpha_0 E(y_{t-1}^2) + \alpha_1 E(y_{t-1}^4) - \frac{\alpha_0^2}{(1-\alpha_1)^2}\right) \sigma_\varepsilon^2 \end{aligned}$$

Ec. 3

De la expresión de la Ec. 3, el único valor no calculado hasta el momento es el siguiente:

$$E(y_{t-1}^4) = E(\varepsilon_t^4 (\alpha_0 + \alpha_1 y_{t-1}^2)^2) = \left( \frac{\alpha_0^2}{(1-\alpha_1)^2} \frac{3(1-\alpha_1)^2}{1-3\alpha_1^2} \right) (\sigma_\varepsilon^2)^2 \quad \text{Ec. 4}$$

Para este último cálculo se suponen dos cuestiones:

1. El ruido blanco se distribuye como una normal, por lo que podemos determinar el cuadrado de la varianza.
2. Además,  $y_t$  es un proceso estocástico estacionario; es decir, el parámetro  $\alpha$  es menor que la unidad y, con ello, nos aseguramos de que  $3\alpha_1^2 < 1$ , y la expresión anterior es siempre calculable.

Sin pérdida de generalidad, todo el proceso se puede realizar con la hipótesis de que la perturbación aleatoria tiene desviación típica igual a uno.

Aplicando este resultado a la ecuación anterior y simplificando, podemos calcular la autocovarianza de orden uno del proceso  $y_t$  al cuadrado como:

$$\gamma_2(1) = \frac{2\alpha_0^2}{(1-\alpha_1)^2(1-3\alpha_1^2)} \neq 0 \quad \text{Ec. 5}$$

Con lo cual, el proceso  $y_t$  descrito, si bien no tiene autocorrelación en forma lineal, sí la tiene su forma cuadrática (el resultado de este último cálculo es siempre distinto de cero, luego hay correlación).

La clave de los modelos que vamos a estudiar está en la distinción entre los momentos condicionales y marginales ("incondicionales"). Entre ellos existe una relación que se conoce con el nombre de ley de expectativas iteradas que, usando la expresión de Ruíz (1994), podría definirse como: "la esperanza de la observación  $y_t$ , o



de una función de ella,  $g(y_t)$ , condicional en información disponible en el momento  $t-\tau$  puede calcularse tomando primero la esperanza condicional en información disponible en  $t-1$ , después calculando la esperanza condicional en  $t-2$  y, así, sucesivamente hasta  $t-\tau$ , es decir:

$$E_{t-\tau}(g(y_t)) = E_{t-\tau}(E_{t-\tau+1}(\dots E_{t-1}(g(y_t))))$$

Para calcular la esperanza marginal, se puede dejar que  $\tau$  tienda a infinito.

## 1.2 Interpretación básica de los modelos de heterocedasticidad condicional

Aunque en capítulos posteriores se hará una descripción de cada uno de los distintos modelos de heterocedasticidad condicional mucho más precisa, parece conveniente centrar la idea subyacente a todos ellos de una forma "menos técnica y más clara".

En esta sección nos apoyaremos en un proceso conocido con el nombre de GARCH, dado que este es una forma generalizada que recoge, como caso concreto, el ARCH (q). Nos centraremos en intentar definir de una forma más sencilla lo que se pretende realizar con la especificación GARCH (p,q) como primera aproximación de potenciales utilidades y sentido de toda la exposición ulterior.

En un modelo GARCH (1,1) hay dos ecuaciones:

- Una primera, donde se hace depender a la variable  $y_t$  del valor de su varianza multiplicada por un cierto término aleatorio que es "ruido blanco"

$$\begin{aligned} y_t &= \varepsilon_t h_t \\ y_t &\rightarrow N(0, h_t) \\ \varepsilon_t &\rightarrow N(0, 1) \end{aligned}$$

- Una segunda, alrededor de un valor medio, representado por el término constante  $\alpha_0$ , donde se hace depender el valor actual de la varianza en el período " $t$ " de los valores que esta haya tenido en el momento anterior ( $t-1$ ) y de la fluctuación aleatoria que también se daba en el pasado. En definitiva, podríamos definir los tres términos como:

$$h_t = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta h_{t-1}$$

- Media  $\alpha_0$  : valor de iniciación en torno al cual se producirán ciertas variaciones. También puede entenderse como el valor medio a largo plazo sobre el que se genera la expectativa inmediata a ser modificada por los dos sumandos que después se detallan posteriormente.
- Sumando  $\alpha_1 y_{t-1}^2$  , innovación sobre la volatilidad que se produjo en el período anterior (término ARCH).
- Sumando  $\beta h_{t-1}$  : realización de la varianza en el último período histórico conocido (término GARCH).

El modelo ARCH (1), como simplificación del aquí presentado, sería un GARCH (0,1), donde no se tendría en cuenta la información sobre la última varianza de la endógena calculable; es decir, la varianza del período anterior.

Matemáticamente, podríamos volver a escribir la especificación del modelo GARCH en dos formas distintas que nos permitirían dar un nuevo enfoque al desarrollo que estamos realizando:

1. En primer lugar, si hacemos una serie de sustituciones recursivas en la fórmula del modelo GARCH planteada, podríamos describir la varianza condicional como una media ponderada de todos los residuos al cuadrado del modelo de la siguiente forma:

$$\begin{aligned} h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 + \beta \sigma_{t-1}^2 \\ h_{t-1} &= \alpha_0 + \alpha_1 y_{t-2}^2 + \beta h_{t-2} \\ &\dots\dots\dots \\ h_t &= \alpha_0 + \alpha_1 y_{t-1}^2 + \beta (\alpha_0 + \alpha_1 y_{t-2}^2 + \beta h_{t-2}) = \dots = \\ &= \frac{1}{1-\beta} + \alpha_1 \sum_{j=1}^{\infty} \beta^{j-1} \varepsilon_{t-j}^2 \end{aligned}$$

Con lo cual, tendríamos que la varianza condicional es el resultado de un valor medio constante sumado a una media ponderada decrecientemente de los valores de la varianza muestral en los períodos precedentes.

2. Si representamos el error en el valor esperado de la variable al cuadrado como  $v_t = \varepsilon_t^2 - h_t$ , podríamos reescribir nuevamente el modelo GARCH (1,1) del siguiente modo:

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta)\varepsilon_{t-1}^2 + v_t - \beta v_{t-1}$$

Según esta nueva formulación, con el modelo GARCH (1,1), propuesta por Pantula (1986) estamos representando un proceso ARMA(1,1) heterocedástico para la serie de los errores al cuadrado. Algunos autores expresan que el proceso ARCH (q) no sería más que un proceso de medias móviles con parámetros cambiantes o variables (por ejemplo, Tsay (1987) o Bera (1992)).

### 1.3. Recorrido histórico y principales variantes de los modelos ARCH

A lo largo del tiempo y desde la propuesta inicial de Engle (1982), la extensiva utilización de modelos de heterocedasticidad condicional autorregresiva, sobre todo en el mundo de la finanzas, ha dado lugar a diversas formulaciones en la definición y evolución de la varianza del proceso estocástico. Las principales preocupaciones en esta materia se han centrado en los siguientes aspectos:

- Dado el carácter positivo de la varianza por construcción estadística, es fundamental que los parámetros estimados sean positivos, por lo que asegurar tanto su posibilidad de cálculo (existencia de los momentos de segundo orden), como su restricción intrínseca (mayores que cero) ha centrado gran parte de los estudios complementarios en la "familia de los ARCH".
- Las particularidades de los mercados financieros y cambiarios han determinado búsquedas de modelos que sean capaces de reproducir no sólo la "volatilidad cambiante", sino también los posibles "efectos contagio" tan comunes en variables financieras. Se ha contrastado empíricamente el "hecho estilizado" de que, en este tipo de mercados, movimientos de fuerte intensidad suelen venir acompañados de otros de igual naturaleza hasta que, en un momento dado, la intensidad se reduce durante períodos consecutivos más o menos largos, a su vez. Reproducir estadísticamente este hecho, exige funciones no lineales, como ya lo es el modelo ARCH seminal, pero, con frecuencia, también exige funciones más complejas.

(exponenciales, logarítmicas, etc.). Determinar modelos capaces de recoger estos frecuentes cambios ha centrado varias aportaciones al desarrollo de los modelos ARCH.

- Desde las primeras aplicaciones realizadas por econométricos diversos, lo que Engle bautizaba con el nombre de "shocks asimétricos en la varianza" ha preocupado en las investigaciones sobre los modelos ARCH. Es frecuente contrastar en las variables evoluciones de períodos al alza menos aceleradas que en períodos de caída, en cuanto a la evolución de las variables. Los modelos definidos como de "Régimen Cambiante" o de "Umbral" (*Threshold ARCH*, *Switching ARCH*, ...) plantean la complejidad estadística de este tipo de estructuras y las formas más "exactas" de realizar estimaciones sobre ellos.

Otra vertiente natural de investigación en el campo de los modelos ARCH concierne a la integrabilidad de los procesos descritos. Con frecuencia, los parámetros que corresponden al proceso de la varianza condicional suman valores muy próximos a uno, haciendo que dejen de existir momentos superiores de este momento (que sean infinitos). En las formulaciones tipo IGARCH y de Memoria Larga que se expondrán posteriormente se hace un detallado análisis del problema conocido como "persistencia en volatilidad".

Planteado un proceso estocástico estacionario en sentido débil y condicionado a la información ya existente sobre su pasado del tipo:

$$\varepsilon_t | \psi_{t-1} \Rightarrow N(0, h_t)$$

En el siguiente cuadro se muestran, de un modo esquemático, las principales aportaciones que se han venido realizando en la ecuación de formación de la varianza condicional heterocedástica a lo largo de estos diecisiete años de existencia de los modelos ARCH. En posteriores capítulos, se hará un pormenorizado análisis de la mayor parte de estas variantes.

Cuadro 2: Principales especificaciones de la "familia arch" a lo largo del tiempo

Año	Nombre	Autor-es	Especificación de la varianza	Aportación principal
1982	ARCH	Engle	$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$	Primera especificación y desarrollo.
1983	Modelos ARCH Multivar.	Kraft y Engle	$H_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 H_{t-1}$ $\varepsilon_t = y_t - xb$	Incorporación de más variables explicativas y desarrollo de los modelos aplicando la matriz de varianzas-covarianzas ( $H_t$ ).
1986	ARCH-M	Engle, Lilien y Robins	$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$	Modelo ARCH incorporando la desviación típica heterocedástica modelizada como explicativa de la media
1986	GARCH y GARCH en Media	Bollerslev	$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1}$	Método generalizado sin restricciones para la estimación de los parámetros ARCH con infinitos retardos.
1986	LGARCH	Bollerslev y Taylor	$h_t = \alpha_0 + \alpha_1 h_{t-1} + \alpha_2 h_{t-1} \varepsilon_{t-1}^2$	Linealización del modelo GARCH-M
1986	MGARCH	Geweke y Pantula	$\ln(h_t) = \alpha_0 + \alpha_1 \ln(\varepsilon_{t-1}^2) + \alpha_2 \ln(h_{t-1})$	Especificación de la varianza multiplicativa (linealizada con logaritmos)
1986	IGARCH	Engle y Bollerslev	$h_t = \alpha \varepsilon_{t-1}^2 + (1 - \alpha) h_{t-1}$	Persistencia en varianza condicional heterocedástica. Modelos integrados en varianza.
1989	EGARCH	Nelson	$\log(h_t) = \alpha_0 + \beta_1 \log(h_{t-1})$ $+ \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \alpha \left[ \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} - \sqrt{2\pi} \right]$	Modelos ARCH para procesos no normales (funciones de densidad exponenciales). Carácter asimétrico de la respuesta a shocks positivos o negativos.
1989	TS-GARCH	Schwert	$\sqrt{h_t} = \alpha_0 + \alpha_1 \sqrt{h_{t-1}} + \alpha_2 \sqrt{h_{t-1}}  \varepsilon_{t-1}^2 $	Corrección de efectos asimétricos en las variaciones al alza y a la baja
1990	AGARCH NGARCH	Engle y Ng	$h_t = \alpha_0 + \alpha_1 h_{t-1} + \alpha_2 h_{t-1} (\varepsilon_{t-1} - c)^2$	Contraste y solución de autocorrelación entre la perturbación aleatoria y su varianza.
1990	FACTOR ARCH	Engle, Ng y Rothschild	$H_t = \sum_{k=1}^K \beta_k \beta_k' \lambda_{kt} + \Omega$	Empleo de la covarianza entre varias series temporales como explicativa de la varianza condicional heterocedástica
1992	T-GARCH	Gourieroux Zakonian (1994)	$\sqrt{h_t} = \alpha_0 + \alpha_1 \sqrt{h_{t-1}} + \alpha_2 \sqrt{h_{t-1}}  \varepsilon_{t-1}^2  + \alpha \sqrt{h_{t-1}} \max(0, \varepsilon_{t-1})^2$	Modelos dinámicos donde media y varianza condicionales son funciones stepwise endógenas
1993	GJR-GARCH	Glosten y Otros	$h_t = \alpha_0 + \alpha_1 h_{t-1} + \alpha_2 h_{t-1} \varepsilon_{t-1}^2 + \alpha h_{t-1} \max(0, \varepsilon_{t-1})^2$	Diferenciación del parámetro en subida y en bajada
1993	V-GARCH	Engle y Ng	$h_t = \alpha_0 + \alpha_1 h_{t-1} + \alpha_2 (\varepsilon_{t-1} / \sqrt{h_{t-1}} + c)^2$	Similar al NGARCH, con una variación mayor en los parámetros asimétricos.

1993	A-PARCH	Ding y otros	$h_t^\delta = \alpha_0 + \sum_{i=1}^p \alpha_i ( \varepsilon_{t-i}  - \gamma_i  \varepsilon_{t-i} )^\delta + \sum_{j=1}^q \beta_j h_{t-j}^\delta$	Se propone modelizar un valor potencial de la desviación típica que atienda al máximo de la función de autocorrelación del valor absoluto del proceso.
1994	Modelos ARCH de Régimen Cambiante	Hamilton y Susmel	$\tilde{\varepsilon}_t = \varepsilon_t / \sqrt{g_{st}}$ $h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \tilde{\varepsilon}_{t-i}^2 + \xi d_{t-1} \tilde{\varepsilon}_{t-1}^2$ $\text{si } \tilde{\varepsilon}_t \leq 0 \quad d_{t-1} = 1$ $\text{resto} \quad d_{t-1} = 0$	Introducción de funciones de densidad que cambian de Normal a t-student a partir de cadenas de Markov. Parámetros ARCH cambiantes a partir de una matriz de "estado" o "régimen" de la variable en el período previo.
1996	Modelo de memoria larga o "ARCH con componentes"	Ding y Granger	$h_t = w h_{1t} + (1-w) h_{2t}$ $h_{1t} = \alpha_1 \varepsilon_{t-1}^2 + (1-\alpha_1) h_{1t-1}$ $h_{2t} = h_t (1-\alpha_2 - \beta_2) + \alpha_2 \varepsilon_{t-1}^2 + \beta_2 h_{1t-1}$	La varianza se especifica en dos partes: una con efectos importantes de muy corta duración en el tiempo y otra con efectos más discretos, pero persistentes en el tiempo
1997	VAR-GARCH		$h_{q(i)} = x' b + \alpha_i \varepsilon_{t-1,i}^2 + \beta_i h_{t-1,i}$ $i = 1, 2, \dots (\text{n}^\circ \text{ variables del VAR})$	Empleo de un VAR con residuos con heterocedasticidad condicional.

En el terreno de la investigación de los modelos ARCH es fuente de controversia y de desarrollo continuo tanto los sistemas de estimación de este tipo de modelos como los métodos de contraste sobre su estructura y existencia.

El supuesto inicial de normalidad en la distribución de las variables aleatorias se ve quebrantado con frecuencia en las diversas aplicaciones que se realizan empleando el tipo de modelos que nos ocupan. Es fundamental la aportación de Woolridge y Bollerslev (1992) con su estimación de "cuasi-máxima verosimilitud, que permite lograr unos valores de la matriz de varianzas-covarianzas no sólo consistente, sino también eficiente sea cual sea la distribución que rijan la perturbación aleatoria del modelo.

Desde la aparición del artículo de Engle (1982) este hecho ha sido motivo de fuerte controversia y la búsqueda de la forma más correcta de estimar los parámetros de un modelo tipo ARCH ha sido objeto de diversas investigaciones, siendo de destacar las siguientes:

- Estimadores por el Método Generalizado de los Momentos, propuesto por Rich y otros (1991). Estos estimadores manifiestan algunas ventajas evidentes (n requieren un conocimiento exacto de la función de distribución de las variables, su sistema resolutivo es más fácilmente tratable que el tedioso recurso iterativo de la máxima

verosimilitud, el método de Hansen (1982) y Newey (1987) proveen de un contraste de especificación intrínseco a la estimación y extremadamente útil). Por supuesto, el sistema cuenta con la desventaja de ser ineficiente frente al de máxima verosimilitud cuando la verdadera función de distribución es efectivamente una normal.

- Estimación de una función GED y disquisiciones sobre la conveniencia de la t-student propuestas por Nelson (1991) para la determinación de los parámetros. En trabajos como el de Engle y González-Rivera (1991) se demuestra que la ineficiencia relativa de emplear la función normal cuando la verdadera función de distribución es la t-student es muy pequeña.

- Estimación cuasi-máximo verosímil de Bollerslev y Wooldridge (1992), demostrando que para una determinada clase de modelos dinámicos parametrizados por los primeros y segundos momentos, el estimador cuasimáximo verosímil es asintóticamente normal. Los autores realizan una serie de experimentos de Montecarlo sobre los que concluyen la ventaja de este tipo de estimadores (medida por el sesgo cometido) frente a los estimadores máximo verosímiles. La conclusión fundamental es que los estimadores máximo verosímiles aplicados a los modelos ARCH son consistentes aun cuando se la verdadera función de distribución no sea la normal.

- Con frecuencia se han propuesto estimaciones no paramétricas de los modelos tipo ARCH empleando Kernel o con cadenas de Fourier. De ello se hacen eco Pagan y Schwert (1991). El primero plantea una media ponderada en la elección del estimador de la varianza condicional donde, la ponderación, se hace depender de la historia existente sobre el fenómeno; es decir, de sus momentos condicionales en el extremo de equiponderación para todas las observaciones pasadas. La propuesta sobre la matriz de ponderaciones se realiza a partir del Kernel Gaussiano de orden uno multiplicado por la matriz de diferencias de orden determinado de las observaciones hasta el período anterior al considerado en cada momento.

En el caso de las cadenas de Fourier, se define el llamado estimador no paramétrico flexible de Fourier definido por Gallant (1981) donde la varianza condicional se representa por polinomios de bajo orden y términos trigonométricos contruidos a partir de la información disponible en el período  $t$  sobre el pasado de la serie.

- Los estimadores propuestos hasta el momento basan sus resultados en la consistencia como principal objetivo, aunque sus autores suelen hacer hincapié en la

pérdida de eficiencia que en ellos se produce. Centrando la importancia en dicha eficiencia de los estimadores obtenidos, destaca la aportación de Drost y Klaassen (1997). Estos autores basan su propuesta en el teorema LAN<sup>1</sup> para los modelos de series temporales, logrando un método de estimación que no necesita imponer ninguna condición en los momentos de la serie.

- La evidente posibilidad de escribir los modelos como un modelo de parámetros cambiantes es puesta de manifiesto por Juan del Hoyo (1992) como posible estimación del modelo ARCH a partir de la representación del proceso de los estados, poniendo como ejemplo una aplicación del filtro de Kalman del modelo de coeficientes variables implícito en el modelo ARCH(1).

En el mismo terreno de la distribución que rige las perturbaciones aleatorias que se modelizan, son de especial interés las disquisiciones teóricas hechas por Daniel Nelson en diversos artículos sobre la conveniencia de emplear una distribución tipo t-student. Sobre el mismo tema redundan Hamilton y Robert Engle y Victor Ng en diversos papers (todos ellos citados en el cuadro anterior). Engle y González-Ribera (1991) realizan un experimento de Montecarlo para cuantificar la pérdida de eficiencia relativa en la estimación máximo verosímil cuando la verdadera distribución de la perturbación es una t-student o una gamma, definiendo un ratio que muestra una fuerte pérdida en el caso segundo de los casos (distribución gamma) y poco relevante en el primero.

Respecto a la utilidad y empleo de los modelos ARCH, existen, por lo menos, dos recopilaciones relativamente recientes muy significativas sobre ello. Son las de Bollerslev y otros de 1992 y 1994. En ellas se citan más de cuatrocientas aplicaciones diferentes que se han producido en los primeros diez-doce años de existencia de este tipo de modelos. Con frecuencia, las aplicaciones se han centrado en el campo de la economía financiera y, más concretamente, en la aplicación de teorías tipo de valoración de riesgos en la construcción de carteras de inversión a partir de una conveniente modelización de la volatilidad o varianza.

---

<sup>1</sup> LAN (locally Asymptotically Normal) obtenido a partir de la definición de la perturbación aleatoria se puede encontrar una "co-evolución" del ruido y su varianza. El sistema iterativo de resolución se juzga a partir de un ratio de verosimilitud sobre los valores iniciales dados a la varianza.



## **2. MODELOS AUTORREGRESIVOS HETEROCEDÁSTICOS CONDICIONALES: ARCH( $q$ )**

## 2.1. Especificación de un modelo ARCH(q)

Los modelos ARCH tienen su origen en un artículo de Engle (1982), en el que se pretendía obtener una predicción adecuada para la inflación en el Reino Unido, sujeta a fuerte volatilidad y con períodos de especial calma o de especial agitación.

El proceso ARCH(q) viene definido por la siguiente expresión:

$$y_t = \varepsilon_t h_t$$

$$h_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i y_{t-i}^2 \quad \text{Ec. 6}$$

Donde se dan las siguientes restricciones:

1.  $\varepsilon_t$  es un proceso idénticamente distribuido con media cero y desviación típica igual a uno.
2. Los parámetros  $\alpha_0 > 0$  y  $\alpha_i \geq 0$  e  $i=1\dots q$ , y, para cumplirse la condición de estacionariedad en media, la suma de todos los parámetros es menor que la unidad.
3. Si  $\varepsilon_t$  es gaussiano y se distribuye según una normal,  $y_t$  es condicionalmente normal y su varianza es  $h_t^2$ .

El proceso utilizado en la sección anterior como ejemplo era, por tanto, un ARCH(1).

Para éste, ya se demostró que:

Cuadro 3: Momentos condicionales y marginales de una ARCH(q)

	<b>Marginal (incondicional)</b>	<b>Condicional</b>
<b>Esperanza</b>	$E(y_t) = (\varepsilon_t^2 (\alpha_0 + \alpha y_{t-1}^2))^{1/2} =$ $= E(\varepsilon_t) E((\alpha_0 + \alpha y_{t-1}^2)^{1/2}) = 0$	$E_{t-1}(y_t) = (\alpha_0 + \alpha y_{t-1}^2)^{1/2} E_{t-1}(\varepsilon_t) = 0$
<b>Varianza</b>	$E(y_t^2) = E(\varepsilon_t^2 (\alpha_0 + \alpha y_{t-1}^2)) =$ $= \sigma_\varepsilon^2 (\alpha_0 + \alpha E(y_{t-1}^2));$ $\sigma_t^2 = E(y_t^2) = \frac{\alpha_0}{1-\alpha} \sigma_\varepsilon^2$	$E_{t-1}(y_t^2) = (\alpha_0 + \alpha y_{t-1}^2) E_{t-1}(\varepsilon_t^2);$ $\sigma_t^2 = (\alpha_0 + \alpha y_{t-1}^2) \sigma_\varepsilon^2$

- Las esperanzas marginal y condicional son iguales a cero.

- La varianza marginal es constante; mientras que
- la varianza condicional depende de los valores que haya tomado  $y_{t-1}^2$ ; luego no es fija.
- La distribución marginal del proceso ARCH(1) tiene una forma desconocida.

Las demostraciones realizadas para  $q=1$  son generalizables para cualquier valor de "q"

## 2.2. Condición de estacionariedad de un ARCH(q)

El objetivo de este epígrafe será demostrar cuál es la condición necesaria para que la varianza del proceso ARCH(1) tenga un límite finito; es decir, sea calculable para un número indeterminadamente amplio de retardos. Aplicando las condiciones al caso concreto del proceso ARCH(1), se persiguen dos objetivos: por un lado, ilustrar mejor las demostraciones matemáticas y, por otro, concretar esta condición para el proceso más habitualmente utilizado.

Si el proceso  $y_t^2$  tiene media nula, varianza  $h_t$  y se distribuye como una normal, a partir del desarrollo polinómico de Taylor podremos escribir el momento de orden "m" según la siguiente expresión:

$$E(y_t^{2m} | \psi_{t-1}) = h_t^{2m} \prod_{j=1}^m (2j-1) = (\alpha_0 + \alpha_1 y_{t-1}^2)^m \prod_{j=1}^m (2j-1) \quad \text{Ec. 7}$$

La parte anterior al productorio, se podría escribir matricialmente para, por ejemplo,  $m=3, 2$  y  $1$  sucesivamente, como:

$$\begin{bmatrix} (h_t^2)^3 \\ (h_t^2)^2 \\ (h_t^2) \end{bmatrix} = \begin{bmatrix} \alpha_1 & 2\alpha_0\alpha_1^2 & \alpha_0^2\alpha_1 \\ 0 & \alpha_1^2 & 2\alpha_0 \\ 0 & 0 & \alpha_1^3 \end{bmatrix} \begin{bmatrix} y_{t-1}^6 \\ y_{t-1}^4 \\ y_{t-1}^2 \end{bmatrix} + \begin{bmatrix} \alpha_0^3 \\ \alpha_0^2 \\ \alpha_0 \end{bmatrix} \quad \text{Ec. 8}$$

Si escribimos el vector  $z_t' = (y_t^{2m}, y_t^{2(m-1)}, \dots, y_t^2)$ , podríamos resumir la expresión matricial anterior y su esperanza (de  $z_t$  condicionada a la información disponible en sobre "t-1") según la siguiente expresión:

$$E_{t-1}(z_t) = E(z_t | \psi_{t-1}) = b + Az_{t-1} \quad \text{Ec. 9}$$

donde "b" es un vector columna de "m" filas y "A" es una matriz triangular superior "m×m". Obviamente, el número de valores diferentes de la matriz A es "m+(m×m-m)/2", con lo cual la expresión es sólo válida para potencias de "y<sub>t</sub>" menores o iguales al resultado de esa operación.

Dado que el proceso ARCH(1) es muy frecuente, se escriben sus dos primeros momentos siguiendo el sistema propuesto por Engle. Así, las dos expresiones, según la Ec. 7 serían las siguientes:

$$\begin{aligned} E((y_t^2)^2 | \psi_{t-1}) &= (\alpha_0 + \alpha_1 y_{t-1}^2)^2 = 3\alpha_0^2 + 3\alpha_1^2 (y_{t-1}^2)^2 + 6\alpha_0\alpha_1 y_{t-1}^2 \\ E(y_t^2 | \psi_{t-1}) &= \alpha_0 + \alpha_1 y_{t-1}^2 \end{aligned} \quad \text{Ec. 10}$$

Si describimos la matriz  $w_t$  como aquel vector columna que recoge los momentos al cuadrado de los valores del proceso ya al cuadrado (por ejemplo,  $E(w_t) = E(z_t)^2 = E(y_t^2)^2$ ). Es decir, si incluimos en dicha matriz los momentos de primer y segundo orden del proceso cuadrado que centra las discusiones sobre los modelos ARCH que estamos reproduciendo, podríamos escribir matricialmente:

$$\begin{aligned} w_t &= (y_t^4, y_t^2) \\ E(w_t | \psi_{t-1}) &= b + Aw_{t-1} = \begin{pmatrix} 3\alpha_0 \\ \alpha_0 \end{pmatrix} + \begin{pmatrix} 3\alpha_1 & 6\alpha_0\alpha_1 \\ 0 & \alpha_1 \end{pmatrix} \begin{pmatrix} w_{t-1}^2 \\ w_{t-1} \end{pmatrix} \end{aligned} \quad \text{Ec. 11}$$

Por sustitución sucesiva y, teniendo en cuenta que  $w_t = z_t^2$ , la expresión anterior podría escribirse, para la información disponible en  $z_{t-k+1}$ , es decir  $z_{t-k}$ , del siguiente modo:

$$E_{t-1}(z_t) = E(z_t | \psi_{t-1}) = b + Az_{t-1} = b + A(b + Az_{t-2}) = (I_m + A + A^2 + \dots + A^{k-1})b + A^k z_{t-k}$$

Dada la progresión geométrica resultante, es fácil comprobar que esta serie sólo tenderá hacia un número finito en la medida en que todos los autovalores de la matriz A estén dentro del círculo unitario, ya que de otro modo las potencias de sus

elementos serían una progresión explosiva (tenderían a infinito). El límite de la expresión anterior cuando "k" tiende a infinito se puede escribir como:

$$\lim_{k \rightarrow \infty} E(z_t | \psi_{t-k}) = (I - A)^{-1} b \quad \text{Ec. 12}$$

límite en cuya definición no entra la observación a la que nos estemos refiriendo (t), luego tendrá el mismo valor para la varianza marginal (incondicional).

Para el caso del ARCH(1), tomando la matriz A que se obtuvo en la Ec. 9, la expresión anterior sería:

$$E(w_t) = \begin{bmatrix} \frac{3\alpha_0^2}{(1-\alpha_1)^2} \frac{1-\alpha_1^2}{1-3\alpha_1^2} \\ \frac{\alpha_0}{1-\alpha_1} \end{bmatrix} \quad \text{Ec. 13}$$

(2x1)

Por otra parte, dado que la matriz A es triangular superior, los autovalores coinciden con los elementos de su diagonal principal; y, tal como se ha dicho, es condición necesaria y suficiente de estacionariedad del ARCH el que dichos autovalores estén comprendidos dentro del círculo unitario. Teniendo en cuenta como se definieron los momentos condicionales de una normal (Ec. 7), los elementos de la diagonal principal de la matriz A serán:

$$\alpha_1^m \prod_{j=1}^m (2j-1) = \prod_{j=1}^m \alpha_1 (2j-1) = \theta_m \quad \text{Ec. 14}$$

Donde  $\theta_m$  es el resultado de un producto de factores monotónicamente crecientes. En definitiva, este valor sólo convergerá hacia un número en la medida en que  $\alpha_1$ , en valor absoluto, sea menor que uno para el caso del ARCH(1).

Este proceso es generalizable para un ARCH(q). En este caso, se podría decir que dicho proceso, con  $\alpha_0 > 0, \alpha_1, \alpha_2, \dots, \alpha_q \geq 0$ , es estacionario en covarianza si, y sólo si,

todas las raíces asociadas a su función característica están fuera del círculo unitario<sup>2</sup>, lo que es igual que decir que todos los autovalores de la matriz A que resultaría de escribir el sistema para "q+1" parámetros estuvieran dentro del círculo unitario. Para un ARCH(q), la varianza estacionaria del proceso vendrá dada por la expresión:

$$\frac{\alpha_0}{1 - \sum_{j=1}^q \alpha_j} \quad \text{Ec. 15}$$

De la ecuación Ec. 13 se puede inferir que, dado que la varianza marginal del proceso ARCH(1) es  $\text{var}(y_t^2) = \frac{\alpha_0}{1 - \alpha_1}$ , que este proceso tiene unas colas de distribución más de tres veces superiores a las de la normal. Desde la Ecuación Ec. 14 el momento de orden cuatro se puede escribir:

$$E((y_t^2))^2 = \frac{3 \text{var}(y_t^2)}{1 - 3\alpha_1^2}$$

donde, cumpliéndose las condiciones de estacionariedad, el denominador ha de ser menor que uno.

### 2.3. Estimación de un ARCH(q)

En la definición hecha de los procesos de heterocedasticidad condicional autorregresiva se hace el supuesto inicial de que la perturbación aleatoria del modelo propuesto se distribuye según una normal. Engle, en su artículo de 1982, propone la estimación de los parámetros a partir de la maximización del logaritmo de la función de densidad de la normal. La única particularidad del método propuesto, parte de la consideración de la esperanza condicional para la simplificación de los momentos de primero y segundo orden que en ella se obtienen. Obviamente, tal y como viene

<sup>2</sup> Para la demostración, ver Engle, R.F. (1982): *Autorregresive Contional Heterocedasticity with Estimates of The Variance of U.K. Inflation* Econometrica, 50 Pag: 1005-1006, apéndice II.

asumiendo este tipo de modelos, podemos considerar la información del momento "t-1" como conocida en el momento "t".

En el desarrollo matemático-estadístico que se estudia, tanto para el caso del ARCH(q) como para el del modelo de regresión con estructura ARCH, se procede del siguiente modo:

- A partir de la función de densidad de la normal, se explicitan los momentos de primer y segundo orden.
- Se obtiene la matriz de información del suceso conjunto como la media de las esperanzas condicionales de la matriz de momentos de segundo orden (hesiano).
- Se expone el procedimiento de estimación iterativa de los parámetros de los modelos a partir del algoritmo de "scoring"<sup>3</sup>, empleando entonces la matriz de información como sustituta o simplificación del hesiano.
- Se aplica el procedimiento iterativo resultante de la maximización de la expresión del polinomio de Taylor de segundo grado como aproximación de cálculo del valor mínimo de la función verosímil en un punto.

Pasamos ahora a realizar estos pasos de modo detallado.

Siendo  $y_t$  un proceso definido por un ARCH(q) del modo que se presentó en la sección anterior:

---

<sup>3</sup> En este algoritmo se plantea la aproximación del valor de una función en un punto  $\theta_0$  cercano al mínimo de la función  $\theta$  mediante un polinomio de Taylor de segundo grado:

$$F(\theta_0) = M = F(\theta) + [\nabla F(\theta)]'(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)'[\nabla^2 F(\theta)](\theta - \theta_0)$$

El valor de  $\theta_0$  que minimiza la función será el que anula la primera derivada:

$$\frac{\partial M}{\partial \theta_0} = \nabla F(\theta) + [\nabla^2 F(\theta)]^{-1}(F(\hat{\theta}) - \theta_0) = 0$$

$$\hat{\theta}_{n+1} = \hat{\theta} = \hat{\theta}_n - [\nabla^2 F(\theta)]^{-1} \nabla F(\hat{\theta})$$

Con el procedimiento de "scoring" se propone como aproximación sustituir la matriz de derivadas segundas de la maximización verosímil por la matriz de información.

$$y_t = \varepsilon_t h_t$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i y_{t-i}^2 \quad \text{Ec. 16}$$

la media de  $y_t$  es cero y todas sus covarianzas nulas. Además sus observaciones se distribuyen como una normal, dado que es una combinación lineal de una perturbación aleatoria que así se distribuye. Por ello, su distribución conjunta es el producto de las densidades de cada uno de sus elementos, con lo que el logaritmo de esta función será la suma de todas las funciones de densidad.

Podríamos escribir entonces la media del logaritmo de verosimilitud  $l$  en la observación  $t$ , para una muestra de  $T$  observaciones, como:

$$l = \frac{1}{T} \sum_{t=1}^T l_t \quad \text{Ec. 17}$$

El logaritmo de verosimilitud de la normal correspondiente al proceso ARCH(q) antes descrito (al margen de las constantes) sería:

$$l_t = -\frac{1}{2} \log h_t - \frac{1}{2} \frac{y_t^2}{h_t} \quad \text{Ec. 18}$$

La determinación de los parámetros del modelo anterior ( $\alpha_i$   $i=0\dots q$ ), a los que generalizaremos como  $\theta$ , se pueden estimar a partir de la maximización de esta función, siendo las condiciones de primero y segundo orden las siguientes:

$$\frac{\partial l_t}{\partial \theta} = \frac{1}{2h_t} \frac{\partial \sigma_t^2}{\partial \theta} \left( \frac{y_t^2}{h_t} - 1 \right) \quad \text{Ec. 19}$$

y el hesiano:

$$\frac{\partial^2 l_t}{\partial \theta \partial \theta'} = \frac{1}{2(h_t)^2} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial h_t}{\partial \theta'} \left( \frac{y_t^2}{h_t} \right) + \left( \frac{y_t^2}{h_t} - 1 \right) \frac{\partial}{\partial \theta'} \left( \frac{1}{2h_t} \frac{\partial h_t}{\partial \theta} \right) \quad \text{Ec. 20}$$



En la expresión anterior, se pueden simplificar algunos valores teniendo en cuenta el resultado de la esperanza condicional de uno de sus términos:

$$E_{t-1}\left(\frac{y_t^2}{h_t}\right) = \frac{1}{h_t} E_{t-1}(y_t^2) = \frac{h_t}{h_t} = 1 \quad \text{Ec. 21}$$

Esto es posible ya que se ha demostrado que la media del proceso  $y_t$  era nula, que el valor de  $y_{t-1}$  es conocido en el momento  $t$  y que:

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i y_{t-i}^2 \quad \text{Ec. 22}$$

con lo que la esperanza condicional de la varianza -  $E_{t-1}(h_t)$  - es una constante conocida.

Con todo ello, la expresión del hesiano antes escrito quedaría reducida a:

$$\frac{\partial^2 l_t}{\partial \theta \partial \theta'} = \frac{1}{2(h_t)^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta'} \quad \text{Ec. 23}$$

y la matriz de información, es decir la media del hesiano para las  $T$  observaciones totales (supuesta dependencia entre los valores sucesivos del proceso  $y_t$ ), sería:

$$\zeta_{\theta\theta} = \sum_t \frac{1}{2T} E \left[ \frac{1}{(h_t)^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta'} \right] \quad \text{Ec. 24}$$

Esta matriz de información se puede estimar de forma consistente como:

$$\hat{\zeta}_{\theta\theta} = \frac{1}{2T} \sum_t \left[ \frac{1}{(h_t)^2} \frac{\partial h_t}{\partial \theta} \frac{\partial h_t}{\partial \theta'} \right] \quad \text{Ec. 25}$$

Engle propone escribir estas funciones de una forma más sencilla a partir de la siguiente simplificación:

Siendo  $z_t = (1, y_{t-1}^2, y_{t-2}^2, \dots, y_{t-q}^2)$  y  $\alpha' = (\omega, \alpha_1, \alpha_2, \dots, \alpha_q)$  la ecuación de la varianza se puede escribir como

$$h_t = z_t \alpha$$

y el gradiente y la estimada de la matriz de información, respectivamente, como:

$$\frac{\partial l}{\partial \alpha} = \frac{l}{2h_t} z_t \left( \frac{y_t^2}{h_t} - 1 \right)$$

$$\hat{S}_{\alpha\alpha} = \frac{1}{2T} \sum (z_t' z_t / h_t^2)$$

Para obtener el valor estimado de la matriz de parámetros, Engle (1982) propone utilizar el algoritmo de optimización de "scoring" que, para la expresión anterior, sería el siguiente:

$$\begin{aligned} \hat{\theta}_{t+1} &= \hat{\theta}_t + (\tilde{Z}' \tilde{Z})^{-1} \tilde{Z}' W \\ \tilde{z}_t' &= (1, y_{t-1}^2, y_{t-2}^2, \dots, y_{t-q}^2) / h_t \\ \tilde{Z}' &= (\tilde{z}_1', \tilde{z}_2', \dots, \tilde{z}_q') \text{ matriz de } (q+1) \times T \\ w_t &= (y_t^2 / h_t - 1) \\ W' &= (w_1, w_2, \dots, w_q) \text{ vector fila de } 1 \times T \end{aligned}$$

Como puede demostrarse fácilmente, el estimador así obtenido tendrá como matriz de covarianzas la inversa de la matriz de información. Además, este sistema de estimación paramétrica asegura que alcanzamos un mínimo ya que la matriz de información es definida positiva.

Cuando se utiliza este algoritmo de estimación se suele marcar como criterio de convergencia el estadístico G, que se distribuye como una  $\chi^2$  con "q" grados de libertad y que coincide con el valor de la  $R^2$  de la última regresión efectuada en el algoritmo. Si el valor de esta G es mayor que el tabulado, se interpreta que la diferencia entre el valor obtenido y el real es grande, por lo que se sigue iterando.

## 2.4. Modelo ARCH de Regresión

Hasta el momento, se había considerado  $y_t$  como un proceso estocástico sólo vinculado a la evolución de su varianza condicional y de un ruido blanco. La extensión lógica no puede ser otra que realizar una especificación de  $y_t$  en función de más variables explicativas y aplicar el modelo de heterocedasticidad condicional a la especificación de la varianza del proceso. Es decir, aplicar el modelo ARCH a un proceso con media no nula.

La especificación en este caso es la siguiente:

$$\begin{aligned} y_t | \psi_{t-1} &\rightarrow N(x_t \beta, h_t) \\ \varepsilon_t &= y_t - x_t \beta \\ h_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \end{aligned} \quad \text{Ec. 26}$$

En este modelo, las variables explicativas  $x_t$  pueden incluir a la endógena desplazada y la varianza de  $y_t$  se modeliza en función de los valores al cuadrado del término aleatorio.

Para estimar los parámetros de este modelo propuesto por Engle (1982), se puede maximizar el logaritmo de versosimilitud que no sería otro que el de una normal.

Los MCO pueden emplearse como estimadores insesgados y consistentes de los parámetros  $\beta$ , dado que la definición matemática anterior no incluye dependencia entre la matriz de  $x_t$  y la perturbación aleatoria  $\varepsilon_t$ . Ambos están incorrelacionados gracias a la definición dada de las esperanzas condicionales. Sólo en el caso en que las explicativas incluyeran valores de la endógena desplazada, los estimadores MCO serían inconsistentes, ya que el cuadrado de  $x_t$  estaría correlacionado con el cuadrado del error calculado. En cualquier otro caso, los MCO serían insesgados y consistentes, pero no eficientes. Para lograr los estimadores óptimos habría que recurrir a la estimación máximo-verosímil.

Los estimadores máximo-versosímiles se obtienen a partir de la maximización del logaritmo de la función de densidad de la normal, obteniéndose, como condición de primer orden, la siguiente expresión para su cálculo:

$$\begin{aligned}
l_i &= -\frac{1}{2} \log(h_i) - \frac{1}{2} \frac{\varepsilon_i^2}{h_i} \\
l &= \frac{1}{T} \sum_{i=1}^T l_i \\
\frac{\partial l}{\partial \beta} &= \frac{\varepsilon_i x_i}{h_i} + \frac{1}{2h_i} \frac{\partial h_i}{\partial \beta} \left( \frac{\varepsilon_i^2}{h_i} - 1 \right)
\end{aligned}
\tag{Ec. 27}$$

El primer sumando de la expresión anterior es la corrección clásica de heterocedasticidad y el segundo es el resultado de que la varianza también dependa de los parámetros  $\beta$ .

En el caso de un ARCH(1), el resultado de la derivada parcial de la varianza respecto a los parámetros es el siguiente:

$$\frac{\partial h_i}{\partial \beta} = 2(\alpha_1 x_{i-1}^2 \beta - \alpha_1 y_{i-1} x_{i-1}) = -2\alpha_1 \varepsilon_{i-1} x_{i-1} \tag{Ec. 28}$$

Generalizando la expresión para un ARCH(q) y sustituyéndola en la Ec. 28, obtenemos un resultado simplificado de la condición de primer orden:

$$\frac{\partial l}{\partial \beta} = \frac{1}{T} \sum \left[ \frac{\varepsilon_i x_i}{h_i} + \frac{1}{h_i} \left( \frac{\varepsilon_i^2}{h_i} - 1 \right) \sum_j \alpha_j \varepsilon_{i-j} x_{i-j} \right] = \frac{1}{T} \sum \varepsilon_i x_i \left[ h_i^{-1} - \sum_j \alpha_j h_{i+j}^{-2} (\varepsilon_{i-j} - h_{i+j}) \right]$$

La condición de segundo orden, es decir, el hesiano, quedaría de la siguiente forma:

$$\frac{\partial^2 l}{\partial \beta \partial \beta'} = -\frac{x_i' x_i}{h_i} - \frac{1}{2h_i^2} \frac{\partial h_i}{\partial \beta} \frac{\partial h_i}{\partial \beta'} \frac{\varepsilon_i^2}{h_i} - 2 \frac{\varepsilon_i x_i}{h_i^2} \frac{\partial h_i}{\partial \beta} + \left( \frac{\varepsilon_i^2}{h_i} - 1 \right) \frac{\partial}{\partial \beta'} \left[ \frac{1}{2h_i} \frac{\partial h_i}{\partial \beta} \right]$$

En el hesiano, tomando esperanzas condicionales, la derivada de la varianza  $h_i$  respecto a los parámetros es nula, ya que esta depende de términos del pasado perfectamente conocidos. Por otro lado, el cociente del residuo al cuadrado sobre la varianza se hace igual a uno.

Dado que el proceso ARCH supone dependencia entre las distribuciones de probabilidad del proceso en sucesivos momentos del tiempo, debe definirse la matriz

de información por el valor esperado de las esperanzas de la varianza condicional, lo que, en este caso, resultaría:

$$\zeta_{\beta\beta} = \frac{1}{T} \sum_i E \left[ E \left( \frac{\partial^2 l}{\partial \beta \partial \beta'} \middle| \psi_{t-1} \right) \right] = \frac{1}{T} \sum_i E \left[ \frac{x_i' x_i}{h_i} - \frac{1}{2h_i^2} \frac{\partial h_i}{\partial \beta} \frac{\partial h_i}{\partial \beta'} \right]$$

Esta matriz de información se puede estimar consistentemente como:

$$\hat{\zeta}_{\beta\beta} = \frac{1}{T} \sum_i \left[ \frac{x_i' x_i}{h_i} - 2 \sum_j \alpha_j^2 \frac{\varepsilon_{t-j}^{-2}}{h_i^2} x_{t-j}' x_{t-j} \right] \quad \text{Ec. 29}$$

En su artículo pionero, Engle (1982) demuestra que, si el proceso ARCH(q) es simétrico, la estimación independiente de los parámetros del modelo y de su estructura ARCH es asintóticamente eficiente ya que la matriz de información cruzada de los parámetros es nula ( $\zeta_{\alpha\beta} = 0$ ), proponiéndose los siguientes pasos para la estimación:

- Estimar la matriz de parámetros  $\beta$  por MCO:

$$\bullet \quad \hat{\beta}^{MCO} = [x_i' x_i]^{-1} x_i' y_i$$

- Utilizar los valores del error obtenidos como estimaciones de  $\varepsilon_t$ :

$$\bullet \quad \hat{\varepsilon}_t = y_t - x_t' \hat{\beta}$$

- Con los valores del error al cuadrado, determinar los valores de los parámetros  $\alpha$  a partir del algoritmo iterativo de "scoring" de Máxima verosimilitud.

$$\hat{\alpha}^{i+1} = \hat{\alpha}^i + [\hat{\zeta}_{\alpha\alpha}^i]^{-1} \frac{1}{T} \sum_i \frac{\partial l_i^i}{\partial \alpha}$$

Para el ARCH(q), la expresión anterior se resume en:

$$\begin{aligned}\hat{\alpha}^{i+1} &= \hat{\alpha}^i + (\tilde{z}_i' \tilde{z}_i)^{-1} \tilde{z}_i' f^i \\ \tilde{z}_i &= (1, \hat{\varepsilon}_{i-1}^2, \hat{\varepsilon}_{i-2}^2, \dots, \hat{\varepsilon}_{i-q}^2) / h_i^i \\ \tilde{z}_i' &= (1, \tilde{z}_1', \tilde{z}_2', \dots, \tilde{z}_T') \\ f_i^i &= (\hat{\varepsilon}_{i-2}^2 - h_i^i) / h_i^i \\ f^i &= (f_1^i, f_2^i, \dots, f_T^i)\end{aligned}$$

En donde  $\hat{\varepsilon}_i$  es el residuo,  $h_i^i$  es la varianza estimada y  $\alpha^i$  es el vector de parámetros, todos ellos en la iteración "i".

El criterio de convergencia se puede fijar, según Belsley (1977) y adaptado al modelo que estamos contemplando, en:

$$R^2 = \frac{\partial l'}{\partial \alpha} \left( \frac{\partial l'}{\partial \alpha \partial \alpha'} \right)^{-1} \frac{\partial l}{\partial \alpha}$$

Tal que  $R^2$  es el coeficiente de correlación cuadrado de la regresión auxiliar.

- Reproducir todo el proceso cuando se haya obtenido un valor de la fase (3) hasta el valor de convergencia señalado.

Los valores de las nuevas  $\alpha$  estimadas a partir de la información de la matriz de varianzas-covarianzas correctamente calculada en la fase (3) se podría aproximar nuevamente con un "scoring" del siguiente modo:

$$\hat{\alpha}^{i+1} = \hat{\alpha}^i + \left[ \hat{\zeta}_{\alpha\alpha} \right]^{-1} \frac{\partial l_i}{\partial \alpha}$$

Donde se puede emplear como matriz de información estimada la proveniente de la ecuación antes deducida.

## 2.5. Contrastes de estructura ARCH

Dada que la estimación óptima en presencia de heterocedasticidad condicional es la máximo versosímil, Engle propone como sistema de contrastación previo al desarrollo del modelo ARCH el contraste: del multiplicador de Lagrange.

El estadístico propuesto es el cociente entre la hipótesis nula del modelo restringido ( $\alpha_1=\alpha_2=\dots=\alpha_q=0$ ) frente al modelo de la varianza  $h_t = h(z_t, \alpha)$  donde  $h$  es una función derivable y  $z_t = (1, \hat{\varepsilon}_{t-1}^2, \hat{\varepsilon}_{t-2}^2, \dots, \hat{\varepsilon}_{t-q}^2)$ , recogiendo los valores de los residuos del modelo estimado por MCO. Bajo la hipótesis nula,  $h_t$  es una constante. Engle propone la aproximación al multiplicador de Lagrange de Breusch (1978) y Godfrey (1978)<sup>4</sup>, con lo que la expresión para este test sería:

$$\zeta = T f^{0'} z (z' z)^{-1} z' f^0 / f^{0'} f^0 = T \cdot R^2$$

$$T \cdot R^2 \rightarrow \chi_q^2$$

Cuando el valor del estadístico experimental supere al tabulado, se rechazará la hipótesis nula, por lo que se puede considerar la existencia de una estructura heterocedástica condicional en los residuos del modelo.

<sup>4</sup> BREUSCH, T.S. y PAGAN, A.R. (1978): *A Simple Test for Heterocedasticity and Random Coefficient Variance*. *Econometrica*, 46 pgs. 1287-1294.

GODFREY, L.G. (1978): *Testing Against General Autoregressive and Moving Average Models when the Regressors include Lagged Dependent Variables*. *Econometrica*, 46, pgs. 1293-1302.

Estos autores proponen la estimación del modelo, mediante mínimos cuadrados ordinarios, dos veces, una para obtener los errores y otra segunda del modelo inicial más una especificación accesoria en función de los errores retardados obtenidos:

$$1. y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \beta_{k+1} y_{t-1} + \beta_{k+2} y_{t-2} + \dots + U_t \text{ donde } e_t = y_t - \hat{y}_{t-1}, \text{ y}$$

2.

$$y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \beta_{k+1} y_{t-1} + \beta_{k+2} y_{t-2} + \dots + \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \dots + \varphi_p e_{t-p} + U_t$$

Los autores demuestran que la expresión  $T \cdot R^2$  ( $T$ : tamaño de la muestra y  $R^2$ : coeficiente de correlación lineal de la regresión) se distribuye asintóticamente como una  $\chi^2$  con "p" grados de libertad. Se rechazará la hipótesis nula de no existencia de autocorrelación si esta expresión es superior al valor de tablas

De modo alternativo, Engle y Kraft (1983) proponen un contraste para determinar si el resultado de la aplicación del ARCH produce una varianza condicional homocedástica o no. La expresión del contraste es la siguiente:

$$h_t \alpha_0 + \alpha_1 \sum_{i=1}^q \frac{(q+1-i)}{36 \varepsilon_{t-i}^2}$$

En definitiva, para verificar la presencia o no de un proceso ARCH(q) a partir del Multiplicador de Lagrange, Engle propone que se sigan los siguientes pasos:

1. Emplear mínimos cuadrados ordinarios en la estimación de un modelo autorregresivo del tipo:

$$z_t = a_0 + a_1 z_{t-1} + a_2 z_{t-2} + \dots + a_k z_{t-k} + \varepsilon_t$$

2. Obtener la serie de los errores al cuadrado y estimar la siguiente regresión:

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2$$

Si no estamos ante un proceso ARCH, el valor de los parámetros obtenidos en la regresión del paso 2 debieran ser nulos. A partir de una muestra suficientemente grande de datos, el test  $NR^2$  converge a una  $\chi^2$  con "q" grados de libertad. Si el valor de este contraste es suficientemente grande (comparado con el valor de tablas), estaríamos rechazando la hipótesis nula, o, lo que es igual, aceptaríamos la existencia de parámetros distintos de cero y, por tanto, de un modelo con varianza heterocedástica condicional; es decir, un ARCH.

## 2.6. Predicción con modelos ARCH

En un modelo autorregresivo de orden uno sobre un proceso  $y_t$  estacionario en desviaciones a la media escrito como:



$$y_t = \phi y_{t-1} + \varepsilon_t$$

$$\varepsilon_t \Rightarrow N(0; \sigma_\varepsilon^2)$$

Ec. 30

se puede escribir el valor para el período extramuestral "t+s" y calcularse, posteriormente, los primeros momentos marginales y condicionales como sigue:

Cuadro 4: Momentos marginales y condicionales del proceso AR(1) para la observación "t+s"

$y_{t+s} = \phi y_{t+s-1} + \varepsilon_{t+s} = \phi(\phi y_{t+s-2} + \varepsilon_{t+s-1}) + \varepsilon_{t+s} = \dots =$ $= \phi^s y_t + \sum_{i=0}^s \phi^{s-i} \varepsilon_{t+s-i}$		
Momentos	Marginales	Condicionales
Esperanza	$E(y_{t+s}) = E(\phi y_{t+s-1} + \varepsilon_{t+s}) = 0$	$E_t(y_{t+s}) = \phi^s E_t(y_t) + \sum_{i=0}^s \phi^{s-i} E_t(\varepsilon_{t+s-i}) =$ $= \phi^s y_t$
Varianza	$E(y_{t+s} - E(y_{t+s}))^2 = E(\phi y_{t+s-1} + \varepsilon_{t+s})^2 =$ $\phi^2 E(y_{t+s-1})^2 + E(\varepsilon_{t+s})^2 = \sigma_\varepsilon^2 (1 - \phi^2)^{-1}$	$E_t(y_{t+s} - E_t(y_{t+s}))^2 =$ $= E_t(\phi^s y_t + \sum_{i=0}^s \phi^{s-i} \varepsilon_{t+s-i} - \phi^s y_t)^2 = \sigma_\varepsilon^2 \sum_{i=0}^s \phi^{s-i}$

Nótese que, en la expresión del valor de predicción para el período "t+s" utilizada en el cuadro anterior:

$$y_{t+s} = \phi^s y_t + \sum_{i=0}^s \phi^i \varepsilon_{t+s-i}$$

↓

↓

↓

Predicción

Esperanza condicional

Error de Predicción

Con esta formulación, los valores de predicción dependen del horizonte temporal elegido (s), pero no la información disponible hasta el período actual "t" de la muestra histórica.

De la forma en la que ya ha sido comentado en secciones anteriores, en un proceso ARCH(1) describimos la varianza condicional de la perturbación aleatoria también como un proceso autorregresivo:

$$\begin{aligned} y_t &= \phi y_{t-1} + \varepsilon_t \\ \varepsilon_t &\Rightarrow N(0; \alpha_0(1 - \alpha_1^2)^{-1}) \\ E_{t-1}(\varepsilon_t) &= 0 \\ V_{t-1}(\varepsilon_t) &= h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \end{aligned} \quad \text{Ec. 31}$$

donde la perturbación aleatoria está incorrelacionada serialmente, aunque existe dependencia en sus segundos momentos.

La varianza marginal de la perturbación seguiría siendo la misma que para un AR(1):

$$Var(\varepsilon_t) = \sigma_\varepsilon^2 = \alpha_0(1 - \alpha_1)^{-1} \quad \text{Ec. 32}$$

de donde, para poder comparar entre las varianzas condicional y marginal, se puede despejar el término independiente como  $\sigma_\varepsilon^2(1 - \alpha_1) = \alpha_0$  y sustituirlo en la Ec. 31, de lo cual resultaría:

$$\begin{aligned} h_t &= \sigma_\varepsilon^2(1 - \alpha_1) + \alpha_1 \varepsilon_{t-1}^2 \\ h_t - \sigma_\varepsilon^2 &= \alpha_1(\varepsilon_{t-1}^2 - \sigma_\varepsilon^2) \end{aligned} \quad \text{Ec. 33}$$

De la Ec. 33 se pueden extraer las siguientes conclusiones:

- La varianza condicional será superior a la marginal en la medida en que el cuadrado de la perturbación en el período previo así lo sea.
- Si la perturbación se distribuye como una normal, el cuarto momento marginal ( $E(\varepsilon_t^2)^2$ ) será  $3\alpha_1^4$  veces mayor que el momento marginal del mismo orden<sup>5</sup>, mostrando entonces el proceso una cola de distribución mucho más larga (es leptocúrtico).

De cara a obtener el valor de predicción en el caso planteado en la Ec. 31, nuevamente podemos escribir los dos primeros momentos condicionales como:

$$E_t(y_{t+1}) = \phi y_t \quad \text{Ec. 34}$$

$$Var_t(y_{t+1}) = E_t(\phi y_t + \varepsilon_{t+1} - \phi y_t)^2 = E_t(\varepsilon_{t+1})^2 = h_{t+1} = \alpha_0 + \alpha_1(y_t + \phi y_{t-1})^2 \quad \text{Ec. 35}$$

Ahora, tanto la esperanza como la varianza dependen del conjunto de información disponible en el momento "t". De hecho, la varianza condicional es tanto mayor cuanto mayor sea el valor de la perturbación en el momento anterior ya conocida  $E_t(\varepsilon_{t+1})^2$  en "t+1".

Esta situación se puede generalizar para el valor de predicción en "t+s" a partir de sustituciones sucesivas:

$$Var_t(y_{t+s}) = \sum_{i=1}^s \phi^{2(s-i)} E(h_{t+i}) \quad \text{Ec. 36}$$

uniendo al resultado de la Ec. 33 la ley de expectativas iteradas, podemos escribir:

$$E(h_{t+i}) = \sigma_\varepsilon^2 + \alpha_1 E(h_{t+i-1} - \sigma_\varepsilon^2) \quad \text{Ec. 37}$$

luego la varianza condicional para el período "t+s" de la Ec. 36 se puede escribir como:

$$Var_t(y_{t+s}) = \sigma_\varepsilon^2 \sum_{i=0}^{s-1} \phi^{2i} + \alpha_1^{s-1} (h_{t+1} - \sigma_\varepsilon^2) \sum_{i=0}^{s-1} \phi^{2i-1} \alpha_1^{-i} \quad \text{Ec. 38}$$

Como resultado de esta demostración se pueden concluir dos evidencias matemáticas:

---

<sup>5</sup> En la sección dedicada a la estacionariedad del proceso ARCH(q) se dedujo esta diferencia.

- En el modelo ARCH(q) las predicciones de la varianza sí están sujetas a los valores ya existentes sobre información del pasado. Esto es obvio a partir del segundo sumando de la Ec. 38, en el cual está presente la varianza  $h_{t+1}$ , que viene, a su vez y según la estructura ARCH, definida por el valor de  $h_t$  modificado por  $\alpha_1$ .
- Las predicciones a largo plazo darán lugar a una varianza de predicción tanto más parecida a la del modelo autorregresivo AR(1) cuanto menor sea el valor del parámetro del ARCH(1) expuesto; es decir, cuanto más pequeño sea  $\alpha_1$  y más a largo plazo sea la predicción, la expresión de la Ec. 38 más parecida será a la varianza condicional recogida en el Cuadro 4, ya que el segundo sumando tendería a eliminarse.

De cara a la predicción con modelos ARCH, habrá que calcular los intervalos de predicción correctamente. Con este fin, podemos obtener el error cuadrático medio de predicción para cualquier momento, por ejemplo  $t+2$ , con el siguiente procedimiento:

$$\begin{aligned}
 ECM(\hat{y}_{T+2}) &= E_T(\hat{y}_{T+2} - y_{T+2})^2 = E_T(y_{T+2})^2 = E_T(E_{T+1}(y_{T+2}^2)) = \\
 &= \omega + \alpha_2 y_T + \dots + \alpha_q y_{T+1-q} + \alpha_1 E_T(y_{T+1}^2) = \\
 &= \omega(1 + \alpha_1) + (\alpha_1^2 + \alpha_2)y_T^2 + (\alpha_1\alpha_2 + \alpha_3)y_{T-1}^2 + \dots + \alpha_1\alpha_q y_{T+2-q}^2 + \alpha_q y_{T+1-q}^2
 \end{aligned}$$

## 2.7. Implicaciones de un arch(1) sobre la estructura de un modelo

Suponiendo que tenemos caracterizado un proceso como un autorregresivo de orden uno:

$$z_t = a_0 + a_1 z_{t-1} + \varepsilon_t$$

En dicho proceso, se pueden obtener sencillamente media y varianza marginales a partir de la descomposición de Wald, quedando:

### Media

$$\begin{aligned}
 E(z_t) &= E(a_0 + a_1 z_{t-1} + \varepsilon_t) = a_0 + a_1 E(z_{t-1}) + E(\varepsilon_t) = a_0 + a_1 E(z_{t-1}) = \\
 &= a_0 + a_1 E(a_0 + a_1 z_{t-2} + \varepsilon_{t-1}) = \dots = a_0(1 + a_1 + a_1^2 + a_1^3 + \dots) = \frac{a_0}{1 - a_1}
 \end{aligned}$$

Varianza (para simplificar, se realiza sobre el modelo AR(1) en desviaciones a la media, sin que ello suponga más que un cambio de origen de los datos ante el cual, la varianza es insensible):

$$\begin{aligned} E(\tilde{z}_t)^2 &= E(a_1 \tilde{z}_{t-1} + \varepsilon_t)^2 = a_1^2 E(\tilde{z}_{t-1})^2 + E(\varepsilon_t)^2 = (a_1^2 (E(\tilde{z}_{t-2})^2 + E(\varepsilon_{t-1})^2) + E(\varepsilon_t)^2 = \\ &= (a_1^2 (a_1^2 (a_1^2 E(\tilde{z}_{t-3})^2 + E(\varepsilon_{t-2})^2) + E(\varepsilon_{t-1})^2) + E(\varepsilon_t)^2 = \dots = E(\varepsilon_t)^2 (1 + a_1^2 + a_1^4 + a_1^6 + \dots) = \\ &= \frac{E(\varepsilon_t)^2}{1 - a_1^2} \end{aligned}$$

donde  $\tilde{z}_t = z_t - E(z_t)$

Podríamos plantear que el término aleatorio  $\varepsilon_t$  goza de las características que nos han impulsado a hablar de los procesos ARCH; es decir, observando sus primeros momentos marginales, la media de este proceso es nula y su varianza es constante. Sin embargo, observando los momentos marginales veríamos que, si bien la esperanza condicional sigue siendo nula, la varianza es un proceso dependiente del valor observado en el periodo anterior.

Podríamos especificar una función de evolución de esta varianza de  $\varepsilon_t$  tal y como hemos definido el proceso ARCH(1):

$$\varepsilon_t = v_t (w + \alpha \varepsilon_{t-1}^2)^{1/2}$$

donde el término  $v_t$  es un ruido blanco con media nula, varianza constante e igual a uno y autocorrelación nula.

Tal y como se definió previamente, los momentos marginales y condicionales de este proceso son:

Media.-

$$E(y_t) = (E(\varepsilon_t^2 (w + \alpha \varepsilon_{t-1}^2)^{1/2}))^{1/2} = E(\varepsilon_t) E((w + \alpha \varepsilon_{t-1}^2)^{1/2}) = 0$$

Varianza.-

$$E(\varepsilon_t^2) = \frac{w}{1 - \alpha} \sigma_v^2 = \frac{w}{1 - \alpha}$$

En definitiva, la varianza del modelo AR(1) original sobre el que estamos trabajando se podría escribir como:

$$\text{var}(z_t) = E(z_t)^2 = \frac{E(\varepsilon_t)^2}{1 - \alpha_1^2} = \frac{\left(\frac{w}{1 - \alpha}\right)}{1 - \alpha_1^2} = w \frac{1 - \alpha_1^2}{1 - \alpha}$$

Con lo cual, el proceso ARCH(1) supone que la variable de predicción  $z_t$  se incrementa cuando lo hacen cualquiera de los parámetros incluidos tanto en su especificación como un AR(1) como en la especificación de su varianza; es decir  $\alpha_1$  y/o  $\alpha$ . Dado que estos parámetros modifican el valor obtenido en el momento anterior, el proceso así construido tiende a "profundizar" el efecto que se produjo en el momento previo, con lo que se recogería que períodos de alta volatilidad vienen acompañados de otros de igual forma y lo mismo para períodos de baja volatilidad.

Lo dicho también podría escribirse del siguiente modo. Si suponemos normalidad en el proceso  $y_t$  formalizado como:

$$y_t | \psi_{t-1} \rightarrow N(0, h_t) \\ h_t = \alpha_0 + \alpha_1 y_{t-1}^2 \quad / 0 < \alpha_0, \alpha_1 < 1$$

Los momentos de orden impar son nulos por simetría y el valor de los parámetros dentro del círculo unitario nos garantiza la estacionariedad en varianza, con lo que podríamos escribir la varianza marginal del proceso como:

$$E(y_t^2) = \frac{\alpha_0}{1 - \alpha_1} = \sigma_y^2 \quad \text{Ec. 39}$$

despejando el término constante de la ecuación ..... e incluyéndolo en la expresión de la varianza condicional, el proceso ARCH(1) se puede escribir como:

$$h_t^2 = \sigma_y^2(1 - \alpha_1) + \alpha_1 y_{t-1}^2 = \sigma_y^2 + \alpha_1 (y_{t-1}^2 - \sigma_y^2)$$

donde, cumpliéndose las condiciones de estacionariedad, la varianza condicional será superior a la marginal siempre y cuando el cuadrado de la observación del período anterior sea mayor que la varianza marginal. Con ello, "sorpresas altas" (positivas o

negativas) en el período anterior, producen varianzas condicionales elevadas en el período siguiente.

$$\begin{array}{ccccc}
 h_t & = & \sigma_y^2 & + & \alpha_1 (y_{t-1}^2 - \sigma_y^2) \\
 \downarrow & & \downarrow & & \downarrow \\
 \boxed{\text{Var. Condicional}} & & \boxed{\text{Var. Marginal}} & & \boxed{y_{t-1}^2 > \sigma_y^2 \rightarrow h_t > \sigma_y^2}
 \end{array}$$

### **3. MODELO ARCH EN MEDIA: ARCH-M**



### 3.1. Especificación

El modelo ARCH en media es una reproducción del esquema del modelo de regresión ARCH añadiendo la desviación típica condicional de los errores como explicativa de la media de un modelo uniecuacional general. Este desarrollo del modelo de Engle (1982) fue introducido por Engle, Lilien y Robins (1986) y se especifica del siguiente modo:

Sea el proceso  $y_t$  con las siguientes características<sup>6</sup>:

$$\begin{aligned} y_t | X_t, \Pi_t &\rightarrow N(X_t \beta + \delta h_t, h_t^2) \\ h_t^2 &= \alpha' W \eta_t + \gamma' Z_t \end{aligned} \quad \text{Ec. 40}$$

donde:

- $Z_t$ : vector de  $(j \times 1)$  variables endógenas desplazadas.
- $X_t$ : vector de  $(k \times 1)$  variables exógenas puras del modelo.
- $\varepsilon_t = y_t - X_t \beta - \delta h_t$ : error en el período  $t$
- $\eta_t = (\varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2)$ : vector de  $q \times 1$  variables
- $W$ : matriz de  $p \times q$  constantes que sirven para introducir restricciones en el número de valores de la anterior que sirven para determinar la varianza de cada  $y_t$ . Para el caso más restringido,  $W_t$  sería la matriz identidad
- $\alpha$ : vector de  $(q \times 1)$  de parámetros en la ecuación de la varianza que modifican los valores cuadrados del error.
- $\gamma$ : vector de  $(j \times 1)$  de parámetros de la varianza
- $\beta$ : vector de  $(k \times 1)$  de parámetros de la media, modificadores de las variables exógenas del modelo.
- $\delta$ : vector de  $(1 \times 1)$  de parámetros de la media, que modifican a la desviación típica del error como explicativa del modelo general

En definitiva, se está escribiendo la forma matricial de un modelo explicado por "k" variables exógenas puras y la desviación típica de la perturbación aleatoria, que, a su

<sup>6</sup> Dado que a partir de ahora se va a utilizar como explicativa del modelo la desviación típica condicional de los errores, los autores del ARCH-M prefieren llamar a la varianza  $h_t^2$  y a la desviación típica simplemente  $h_t$ . Aquí se mantendrá esta nomenclatura en el desarrollo de este apartado.

vez, viene explicada por "q" valores del error al cuadrado desplazados y "j" por desplazamientos de la endógena y.

$$y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \delta h_t + U_t$$

$$h_t^2 = \alpha_0 + \alpha_1 (w_1 \varepsilon_{t-1}^2 + w_2 \varepsilon_{t-2}^2 + \dots + w_q \varepsilon_{t-q}^2) + \gamma_1 y_{t,1} + \gamma_2 y_{t,2} + \dots + \gamma_j y_{t,j}$$

Todos los vectores de parámetros anteriores pueden combinarse en una sola matriz  $\phi = (\alpha', \beta', \gamma', \delta')$  que tendrá orden  $(m \times 1)$ , donde  $m=q+j+k+1$

### 3.2. Estimación

Si contamos con valores iniciales de los datos, la función logaritmo de verosimilitud del modelo nombrado como ARCH-M se puede escribir como (prescindiendo de términos constantes):

$$L(\phi) = \sum L_t(\phi); \quad L_t(\phi) = -\log h_t - \frac{\varepsilon_t^2}{2h_t^2} \quad \text{Ec. 41}$$

Las condiciones de primer orden para el valor máximo verosímil de esta función son:

$$\frac{\partial L_t}{\partial \phi} = \sum ([\varepsilon_t^2 - h_t^2 - h_t \delta \varepsilon_t] h_t^{-4}) \partial h_t^2 / \partial \phi / 2 - \sum [\varepsilon_t / h_t^2] \partial \beta' / \partial \phi \quad \text{Ec. 42}$$

En la ecuación precedente, las derivadas parciales de los parámetros sirven únicamente para identificar cuales están incluidos en una ecuación concreta y cuales no, siendo simplemente un vector de ceros y unos.

Para la estimación de este modelo, habrá que tener en cuenta que la varianza condicional se ha definido en función de innovaciones anteriores y que la ecuación básica también depende de la desviación típica de esta. Por ello, será necesario estimar la derivada parcial de la varianza respecto a los parámetros de forma recursiva  $(\partial h_t^2 / \partial \phi)$ . Los creadores del ARCH-M proponen iniciar este sistema recursivo con la estimación inicial de la Ec. 42 sin tener en cuenta la dependencia de la varianza de los

parámetros. Es decir, maximizando solamente el valor del segundo sumatorio para dar unos valores iniciales del error.

Si se denomina el vector fila de las derivadas del logaritmo de verosimilitud como  $S$ , cada una de sus filas tendrá la expresión:  $[S]_i = \partial L_i / \partial \phi_i$ , pudiendo escribirse la condición de máximo (valores que anulan la primera derivada) como:

$$\partial L_i / \partial \phi = S' i = 0$$

donde  $i$  es un vector de unos de orden  $(T \times 1)$ .

Expresado el hessiano (matriz de información o de segundas derivadas) como la suma de los  $t$  logaritmos de verosimilitud, se puede escribir como:

$$\zeta_i = -E(\partial^2 L_i / \partial \phi \partial \phi')$$

Para el que la matriz de información, media de todos los hessianos, se expresa de forma consistente como:

$$\zeta = E(S' S / T)$$

De cara a la estimación de los parámetros, los autores proponen el sistema iterativo de Berndt y otros (1974), según el cual, el algoritmo de resolución sería:

$$\phi^{i+1} = \phi^i + \lambda (S' S)^{-1} S' i \quad \text{Ec. 43}$$

En el cual,  $S$  es la matriz de primeras derivadas evaluada en el valor del parámetro  $\phi_i$  y  $\lambda$  es el valor o criterio de convergencia elegido que marca el diferencial entre dos resultados de estimación para llegar al resultado final, ajustado inicialmente a la unidad.

Bajo restricciones de regularidad y supuesta una distribución normal, esta expresión cumple la propiedad expresada por Crowder (1976) de:

$$(S' S)^{1/2} (\phi^* - \phi^0) \rightarrow N(0, I) \quad \text{Ec. 44}$$

donde se expresa la diferencia entre el parámetro estimado según la expresión previa y el máximo verosímil.

Los mismos autores del modelo ARCH-M entienden que se puede contrastar el valor obtenido en la estimación de los parámetros a partir del Multiplicador de Lagrange, sobre la ecuación anterior (Ec. 44). El contraste se especificaría según la siguiente ecuación:

$$\Phi_{LM} = i' S^0 (S^{0'} S^0)^{-1} S^{0'} i = T \cdot R_0^2$$

En el cual, la matriz  $S^0$  es la matriz de primeras derivadas evaluada para el parámetro estimado  $\phi^0$ . La  $R_0^2$  será el coeficiente logrado en la regresión en la matriz  $S$  bajo la hipótesis nula de no existencia de proceso ARCH. Este contraste está asintóticamente distribuido como una  $\chi^2$  con tantos grados de libertad como restricciones y cuando la hipótesis nula es cierta. Este contraste se aplica sobre el valor de la primera iteración de la Ec. 43, propuesta para la estimación de los parámetros y que se inicia con el valor de la hipótesis nula.

#### **4. MODELO GENERALIZADO DE HETEROCEDASTICIDAD CONDICIONAL AUTORREGRESIVA: GARCH(p,q)**

#### 4. 1. Especificación

GARCH es la abreviatura de *Generalized Autorregresive Conditional Heteroscedasticity* y da nombre a la ampliación del modelo ARCH ya comentado que realizó Bollerslev (1986) para los órdenes p,q, y Taylor (1986), para el caso específico de los órdenes 1,1.

El modelo ARCH (q) que antes se presentaba, puede mostrar ciertas dificultades de estimación cuando se aplica a estructuras dinámicas en los cuadrados de las series. Por ejemplo, en las series financieras, el número de retardos a utilizar es muy elevado y ello llevaría a un engorroso número de iteraciones para alcanzar una solución al sistema planteado, pudiendo darse el caso de no encontrar nunca una solución. Por ello, el mismo Engle propuso ya en 1983 ciertas restricciones a los parámetros del ARCH(1) que simplificaban su estimación; pero estas no eran un proceso generalizable, por lo que la aportación de Bollerslev es decisiva a la hora de poder dotar de utilidad al modelo presentado por Engle. Como se verá con más detalle posteriormente, el modelo GARCH es al modelo ARMA igual que el ARCH al AR y, siempre y cuando las condiciones de estacionariedad lo permitan, ambos son susceptibles de escribirse como procesos de medias móviles de orden infinito. Por ello, se podría pensar en el modelo GARCH como un ARCH de orden infinito.

El modelo GARCH (p,q) se podría escribir como:

$$\begin{aligned}\varepsilon_t &= \eta_t h_t^{1/2} \\ \varepsilon_t | \psi_{t-1} &\rightarrow N(0, h_t) \\ h_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}\end{aligned}\quad \text{Ec. 45}$$

Con lo cual, el modelo ARCH(q) anterior no sería más que un caso concreto de este (aquel en el que todos los parámetros  $\beta_i$  son igual a cero).

Definidos:

$$\begin{aligned}A &= [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_q] \\ B &= [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_p] \\ \alpha_i(L) \varepsilon_t^2 &= \alpha_i \varepsilon_{t-i}^2\end{aligned}$$

La escritura de la ecuación del modelo para la varianza en función del polinomio de retardos sería:

$$h_t = \alpha_0 + A(L)\varepsilon_t^2 + B(L)h_t \quad \text{Ec. 46}$$

El modelo GARCH (1,1) tiene las siguientes características:

1.  $\eta_t$  es proceso estocástico estacionario idénticamente distribuido con media cero y desviación típica igual a uno.
2. Los parámetros  $\alpha_0 \geq 0$  y  $\alpha_i, \beta_i \geq 0$  e  $i=1\dots q$ , y  $j=1\dots p$ . Además, para cumplirse la condición de estacionariedad en media, la suma de todos los parámetros es menor que la unidad, (como se demostrará posteriormente).
3. La función de distribución marginal no es conocida, pero se pueden calcular los primeros momentos y definir el proceso respecto a su media y a su varianza.

Para el proceso GARCH (1,1), dichos momentos serán:

Cuadro 5: Momentos del modelo ARCH en media

	<b>Marginal (incondicional)</b>	<b>Condicional</b>
<b>Esperanza</b>	$E(\varepsilon_t) = \eta_t (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1})^{1/2} =$ $= E(\eta_t) E((\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1})^{1/2}) = 0$	$E_{t-1}(\varepsilon_t) =$ $= (\alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1})^{1/2} E_{t-1}(\eta_t) = 0$
<b>Varianza</b>	$E(\varepsilon_t^2) = E(h_t) = \frac{\alpha_0}{1 - \alpha_1 - \beta}$	$E_{t-1}(\varepsilon_t^2) = h_t$

- donde  $\eta_t$  es un proceso de "ruido blanco" (entre otras, no hay correlación con su pasado, luego tampoco la hay con el pasado de  $\varepsilon_t$ ).
- El proceso generado  $\varepsilon_t$  es también estacionario
- En los momentos condicionales, en "t", el valor de "t-1" es una realización concreta conocida (no aleatoria)

Bollerslev define el término de varianza condicional introducido en la ecuación como un "mecanismo de aprendizaje adaptativo" de un ARCH infinito. Realmente, el proceso se puede escribir, a partir del polinomio de retardos y despejando el término de la varianza como explicativa de la Ec. 46, como:

$$\begin{aligned} h_t - B(L)h_t &= \alpha_0 + A(L)\varepsilon_t^2 \\ h_t &= \alpha_0(1 - B(1))^{-1} + A(L)(1 - B(L))^{-1} \varepsilon_t^2 = \text{Ec. 47} \\ &= \frac{\alpha_0}{1 - \sum_{i=1}^p \beta_i} + \sum_{i=1}^{\infty} \delta_i \varepsilon_{t-i}^2 \end{aligned}$$

Obviamente, para poder llegar a esta expresión, es necesario que todas las raíces de  $(1-B(z))$  estén fuera del círculo unitario. Es decir, es necesario que el proceso sea estacionario al menos en sentido débil.

#### 4.2. Condición de estacionariedad de un proceso GARCH(p,q)

Si escribimos el proceso estocástico del modelo GARCH(p,q) del siguiente modo:

$$\begin{aligned} \varepsilon_t &= \eta_t h_t^{1/2} \\ \varepsilon_t | \psi_{t-1} &\rightarrow N(0, h_t) \\ \eta_t &\rightarrow N(0,1) \end{aligned} \quad \text{Ec. 48}$$

en esta sección se plantea cuál es la condición que deben cumplir los parámetros del modelo GARCH para asegurarnos la posibilidad de calcular el momento de segundo orden; es decir la  $E(\varepsilon_t^2)$ . Para ello, se plantea el siguiente procedimiento de demostración:

1. Definir la varianza del modelo GARCH,  $h_t$ , de un modo que nos permita calcular los distintos momentos sin tener en cuenta la observación a la que nos estemos refiriendo (simplificación del Milhoj (1984)).
2. Expresar el valor de la  $E(\varepsilon_t^2)$  en esta terminología y determinar su condición de cálculo.



Siguiendo los pasos descritos, para obtener una expresión genérica de los momentos para cualquier observación, empezamos sustituyendo sucesivamente en la expresión del proceso GARCH dada como definición en la Ec. 45, y obtenemos:

Partiendo de  $\varepsilon_t = h_t^{1/2} \eta_t$  y sustituyéndolo por el residuo al cuadrado

$$\begin{aligned}
 h_t &= \alpha_0 + \sum_{i=1}^q \alpha_i \eta_{t-i}^2 h_{t-i} + \sum_{i=1}^p \beta_i h_{t-i} = \\
 &= \alpha_0 + \sum_{j=1}^q \alpha_j \eta_{t-j}^2 \left( \alpha_0 + \sum_{i=1}^q \alpha_i \eta_{t-i-j}^2 h_{t-i-j} + \sum_{i=1}^p \beta_i h_{t-i-j} \right) + \sum_{i=1}^p \beta_i \left( \alpha_0 + \sum_{i=1}^q \alpha_i \eta_{t-i-j}^2 h_{t-i-j} + \sum_{i=1}^p \beta_i h_{t-i-j} \right) = \\
 &\quad \vdots \\
 &= \alpha_0 \sum_{k=0}^{\infty} M(t, k) \quad / \quad M(t, k) = \prod_{i=1}^q \alpha_i^{a_i} \prod_{j=1}^p \beta_j^{b_j} \prod_{i=1}^n \eta_{t-S_i}^2
 \end{aligned}$$

donde:

$$\begin{aligned}
 \sum_{i=1}^q a_i + \sum_{j=1}^p b_j &= k \quad y \quad \sum_{i=1}^q a_i = n \quad y \\
 1 \leq S_1 < S_2 < \dots < S_n &\leq \max(kq, (k-1)q + p)
 \end{aligned}$$

A partir de esta "simplificación" o agregación de términos propuesta por Milhoj (1984), logramos un producto en el cual la resolución de  $M(t, k)$  no depende de "t"; es decir, podemos plantear los sucesivos momentos para la observación "t" sin que la sea necesario emplear este término para ello. Como ejemplo de los primeros momentos podemos escribir:

$$\begin{aligned}
 M(t, 0) &= 1 \\
 M(t, 1) &= \sum_{i=1}^q \alpha_i \eta_{t-i}^2 + \sum_{i=1}^p \beta_i \\
 M(t, 2) &= \sum_{j=1}^q \alpha_j \eta_{t-j}^2 \left( \sum_{i=1}^q \alpha_i \eta_{t-i-j}^2 + \sum_{i=1}^p \beta_i \right) + \sum_{j=1}^p \beta_j \left( \sum_{i=1}^q \alpha_i \eta_{t-i-j}^2 + \sum_{i=1}^p \beta_i \right) \\
 &\quad \vdots \\
 M(t, k+1) &= \sum_{i=1}^q \alpha_i \eta_{t-i}^2 M(t-i, k) + \sum_{i=1}^p \beta_i M(t-i, k)
 \end{aligned} \quad \text{Ec. 49}$$

en definitiva, hemos obtenido la expresión válida para cualquier momento del tiempo u observación:

$$E(M(t, k)) = E(M(s, k)) \quad \forall \quad k, t, s \quad \text{Ec. 50}$$

Teniendo en cuenta los resultados obtenidos para  $M(t,k+1)$  y lo expresado en Ec. 50, así como que  $E(\eta_{t-i}^2) = 1$  (según la definición inicial que hemos dado sobre la varianza del proceso definido en la Ec. 48, es fácil, por sustitución sucesiva de los momentos, calcular la esperanza para  $M(t,k+1)$  como:

$$\begin{aligned} E(M(t, k+1)) &= E\left(\sum_{i=1}^q \alpha_i \eta_{t-i}^2 M(t, k) + \sum_{i=1}^p \beta_i M(t, k)\right) = \left(\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i\right) E(M(t, k)) = \\ &\vdots \\ &= \left(\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i\right)^{k+1} E(M(t, 0)) = \left(\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i\right)^{k+1} \end{aligned}$$

Para determinar el valor de la esperanza del proceso estocástico inicial tendremos ahora:

$$\begin{aligned} E(\varepsilon_t^2) &= \alpha_0 E\left(\sum_{k=0}^{\infty} M(t, k)\right) = \alpha_0 \sum_{k=0}^{\infty} E(M(t, k)) = \\ &= \alpha_0 \left(1 - \sum_{i=1}^q \alpha_i - \sum_{i=1}^p \beta_i\right)^{-1} \end{aligned} \quad \text{Ec. 51}$$

En esta expresión, el proceso sólo convergirá en la medida en que:

$$\left(\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i\right) < 1 \quad \text{Ec. 52}$$

En resumen, la expresión del momento de segundo orden del proceso estocástico definido como un GARCH(p,q) sólo es determinable si la suma de los parámetros (que son siempre positivos, para asegurar que el valor de la varianza así estimada lo sea) es igual a uno.

Si se da esta condición, por simetría se puede inferir que la  $E(\varepsilon_t) = 0$  y  $E(\varepsilon_t \varepsilon_s) = 0 / t \neq s$ .

Dada la enorme extensión del empleo de un modelo GARCH(1,1), conviene enumerar alguna de las ventajas que este supone respecto al ARCH(1) y, en este apartado, es importante resaltar la superioridad del modelo planteado por Bollerslev en cuanto a la

existencia de momentos de órdenes superiores a los que serían calculables con un modelo ARCH.

Tal y como se comentaba en la sección 2.2, Ec. 14, para un proceso ARCH(1) se puede deducir que la condición de estacionariedad es:

$$\begin{aligned} \alpha_1^m \prod_{j=1}^m (2j-1) &< 1 \\ \alpha_1^m a_m &< 1 \\ \alpha_1 &< (a_m)^{-1/m} \end{aligned} \quad \text{Ec. 53}$$

En el proceso GARCH(1,1), si el valor de  $\beta_0$  es igual a cero, nos encontraríamos con un ARCH(1) y si  $\alpha_1 > (a_m)^{-1/m}$ , no existiría el momento de orden "2-emésimo". En el proceso GARCH(1,1), el momento "2-emésimo" se define como<sup>7</sup>:

$$E(\varepsilon_t^{2m}) = a_m \left[ \sum_{n=0}^{m-1} a_n^{-1} E(\varepsilon_t^{2n}) \alpha_0^{m-n} \binom{m}{m-n} \mu(\alpha_1, \beta_1, n) \right] x [1 - \mu(\alpha_1, \beta_1, n)]^{-1} \quad \text{Ec. 54}$$

Entonces, incluso si  $\sum_{i=1}^{\infty} \delta_i = \alpha_1 (1 - \beta_1)^{-1} > (a_m)^{-1/m}$ , el momento de orden 2m podría perfectamente existir en el caso del GARCH (0,1) debido a su memoria más larga.

<sup>7</sup> La demostración de esta expresión se puede encontrar en el artículo de Bollerslev (1986): Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31, Pgs:325-326.

### 4.3. Estimación del Modelo GARCH de regresión

El modelo GARCH(p,q) de regresión se puede definir, en forma esquemática, como:

$$\begin{aligned}\varepsilon_t &= y_t - x_t' b \\ \varepsilon_t | \psi_{t-1} &\Rightarrow N(0, h_t) \\ h_t &= z_t' w \\ z_t' &= (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2; h_{t-1}, \dots, h_{t-p}) \\ w' &= (\alpha_0, \alpha_1, \dots, \alpha_q, \beta_1, \beta_2, \dots, \beta_p)\end{aligned}\quad \text{Ec. 55}$$

Si llamamos al vector de parámetros  $\theta = (b', w')$  tal que  $\theta \in \Theta$ , siendo  $\Theta$  un espacio euclídeo compacto en que existen todos los momentos finitos de  $\varepsilon_t$ . Posteriormente y de cara a la fijación del criterio de convergencia será necesario referirse al verdadero valor de los parámetros, al que llamaremos  $\theta_0$ .

La función del logaritmo máximoverosímil se puede escribir como:

$$\begin{aligned}L_t(\theta) &= T^{-1} \sum_{t=1}^T l_t(\theta) \\ l_t(\theta) &= -\frac{1}{2} \log(h_t) - \frac{1}{2} \varepsilon_t^2 h_t^{-1}\end{aligned}\quad \text{Ec. 56}$$

Derivando respecto al vector w de parámetros, obtenemos las condiciones de primer y segundo orden como:

$$\begin{aligned}\frac{\partial l_t}{\partial w} &= \frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial w} \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \\ \frac{\partial^2 l_t}{\partial w \partial w'} &= \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial w'} \left[ \frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial w} \right] - \frac{1}{2} h_t^{-2} \frac{\partial h_t}{\partial w} \frac{\partial h_t}{\partial w'} \frac{\varepsilon_t^2}{h_t}\end{aligned}\quad \text{Ec. 57}$$

en donde, de forma recursiva y comenzando por obtener una estimación previa para  $h_t$  y  $\varepsilon_t^2, t \leq 0$  que puede ser la varianza muestral, obtenemos la siguiente expresión de la derivada parcial respecto a los parámetros del modelo GARCH:

$$\frac{\partial h_t}{\partial w} = z_t \sum_{i=1}^p \beta_i \frac{\partial h_{t-i}}{\partial w} \quad \text{Ec. 58}$$

Para estimar ahora los parámetros del modelo de regresión, es decir "b", se deriva la función:

$$\begin{aligned} \frac{\partial l_t}{\partial b} &= \varepsilon_t x_t h_t^{-1} + \frac{1}{2} h_t^{-2} \frac{\partial h_t}{\partial b} \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \\ \frac{\partial^2 l_t}{\partial b \partial b'} &= -h_t^{-1} x_t x_t' - \frac{1}{2} h_t^{-2} \frac{\partial h_t}{\partial b} \frac{\partial h_t}{\partial b'} \left( \frac{\varepsilon_t^2}{h_t} \right) - \quad \text{Ec. 59} \\ &\quad - 2 h_t^{-2} \varepsilon_t x_t \frac{\partial h_t}{\partial b} + \left( \frac{\varepsilon_t^2}{h_t} - 1 \right) \frac{\partial}{\partial b'} \left[ \frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial b} \right] \end{aligned}$$

donde, nuevamente por un procedimiento de resolución recursiva, se puede obtener la derivada parcial de la varianza respecto a los parámetros de la regresión como:

$$\frac{\partial h_t}{\partial b} = -2 \sum_{j=1}^q \alpha_j x_{t-j} \varepsilon_{t-j} + \sum_{j=1}^p \beta_j \frac{\partial h_{t-j}}{\partial b} \quad \text{Ec. 60}$$

Se puede demostrar que los elementos de la matriz de información bloque-diagonal correspondientes a "b" son cero, con lo cual,  $w$  puede ser estimada sin pérdida de eficiencia asintótica con una estimación consistente de "b" y viceversa.

De cara a obtener los estimadores de máxima verosimilitud es necesario recurrir a un sistema iterativo. Bollerslev recomienda el método del algoritmo de Berndt, Hall, Hall y Hausman (1974) - BHHH - y no el propuesto por Engle (el de *Scoring*) ya que los dos elementos recursivos que se han señalado en las ecuaciones anteriores hacen que éste sea demasiado complejo y que la convergencia no esté asegurada.

En el algoritmo BHHH, llamaremos  $\theta^{(i)}$  al vector de parámetros obtenido en la iteración  $i$ -ésima. Para calcular el siguiente valor, se aplica la fórmula:

$$\theta^{(i+1)} = \theta^{(i)} + \lambda_i \left( \sum_{t=1}^T \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} \right)^{-1} \sum_{t=1}^T \frac{\partial l_t}{\partial \theta} \quad \text{Ec. 61}$$

donde la derivada del logaritmo de verosimilitud se evalúa para el valor  $\theta^{(i)}$  y  $\lambda_i$  es el valor o criterio de convergencia elegido que marca el diferencial entre dos resultados de estimación para llegar al resultado final. Las iteraciones para  $w^{(i)}$  y para  $b^{(i)}$  se pueden llevar a cabo de forma separada, ya que la matriz de información es bloque diagonal.

A partir del desarrollo de Weis (1982), se demuestra que el vector estimado con este procedimiento,  $\hat{\theta}_T$ , es consistente para  $\theta_0$  y asintóticamente normal, con media  $\theta_0$  y matriz de varianzas covarianzas igual a  $F = E((\partial l_t / \partial \theta)(\partial l_t / \partial \theta'))$  siendo una estimación consistente de esta última expresión la de  $T^{-1} \sum_{t=1}^T (\partial l_t / \partial \theta)(\partial l_t / \partial \theta')^{-1}$ .

#### 4.4. Contraste de estructura GARCH(p,q)

De cara al contraste de estructura GARCH, Bollerslev establece el paralelismo de esta con la metodología Box Jenkins habitual y, por ello, deriva las funciones de autocorrelación total y parcial. Sin embargo, posteriormente se analizará un sistema más estricto y riguroso para este fin, al igual que se hacía para el modelo ARCH(q).

A partir de una sencilla transformación en la escritura del proceso GARCH(p,q) es inmediata la obtención de la función de autocovarianza. Por ello, Bollerslev plantea la siguiente expresión matemáticamente igual a la vista hasta el momento:

$$\begin{aligned} \varepsilon_t &= h_t^{1/2} \eta_t \\ \varepsilon_t^2 &\Rightarrow N(0, h_t) \\ v_t &= (\varepsilon_t^2 - h_t) = (h_t \eta_t - h_t)(\eta_t^2 - 1)h_t \\ \eta_t^2 &\Rightarrow N(0; h_t) \\ \varepsilon_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \varepsilon_{t-j}^2 - \sum_{j=1}^p \beta_j v_{t-j} + v_t \end{aligned} \quad \text{Ec. 62}$$

Definida la función de autocovarianza (simétrica) como:

$$\gamma_n = \gamma_{-n} = E(\varepsilon_t^2; \varepsilon_{t-n}^2) \quad \text{Ec. 63}$$

Uniendo esta expresión y la lograda en la Ec. 62, y teniendo en cuenta que se ha definido  $v_t$  como un proceso estacionario, se puede llegar a la siguiente conclusión:

$$\begin{aligned} \gamma_n &= E(\varepsilon_t^2; \varepsilon_{t-n}^2) = E \left[ (\alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \varepsilon_{t-j}^2 - \sum_{j=1}^p \beta_j v_{t-j} + v_t) \varepsilon_{t-n}^2 \right] = \\ &= \alpha_0 E(\varepsilon_{t-n}^2) + \sum_{i=1}^q \alpha_i E(\varepsilon_{t-i}^2 \varepsilon_{t-n}^2) + \sum_{j=1}^p \beta_j E(\varepsilon_{t-j}^2 \varepsilon_{t-n}^2) - \sum_{j=1}^p \beta_j E(v_{t-j} \varepsilon_{t-n}^2) + \sum_{j=1}^p \beta_j E(v_t \varepsilon_{t-n}^2) = \\ &= 0 + \sum_{i=1}^q \alpha_i \gamma_{n-i} + \sum_{j=1}^p \beta_j \gamma_{n-j} + 0 + 0 \end{aligned}$$

Bollerslev propone resumir el resultado anterior en una única expresión:

$$\begin{aligned} \gamma_n &= \sum_{i=1}^m \varphi_i \gamma_{n-i} \\ \varphi_i &= \alpha_i + \beta_i \quad i = 1, \dots, q \\ \alpha_i &\equiv 0 \quad \text{para } i > q \\ \beta_i &\equiv 0 \quad \text{para } i > p \\ m &= \max(p, q) \end{aligned} \quad \text{Ec. 64}$$

De forma similar al procedimiento de Yule-Walker, a partir de esta expresión, la función de autocorrelación se expresa como:

$$\rho_n = \gamma_n \gamma_0^{-1} = \sum_{i=1}^m \varphi_i \rho_{n-i} \quad n \geq p+1 \quad \text{Ec. 65}$$

A partir de la Ec. 65, se deriva que las primeras "p" autocorrelaciones dependen directamente del valor de los parámetros  $(\alpha_1, \alpha_2, \dots, \alpha_q, \beta_1, \beta_2, \dots, \beta_p)$ .

Para el caso de la función de autocorrelación parcial de un ARCH(q), la expresión anterior se puede plantear para los primeros "k" retardos, resultando  $\phi_{kk}$  la función de autocorrelación parcial que resuelve el conjunto de las "k" ecuaciones de:

$$\rho_n = \gamma_n \gamma_0^{-1} = \sum_{i=1}^m \varphi_{ki} \rho_{n-i} \quad n = 1, \dots, k$$

En esta expresión, para valores superiores a "q" en un proceso ARCH(q) el valor es nulo:

$$\begin{aligned}\phi_{kk} &\neq 0 & k \leq q \\ \phi_{kk} &= 0 & k > q\end{aligned}$$

Para un proceso GARCH(p,q), la función de autocorrelación parcial es, en general no cero, aunque tiende hacia este valor.

Para obtener un valor estimado del correlograma, con frecuencia se siguen los siguientes pasos:

1. Estimar el modelo inicial del cual se pretende observar si su varianza es condicionalmente heterocedástica o no, por ejemplo:

$$z_t = a_0 + a_1 z_{t-1} + a_2 z_{t-2} + \dots + a_k z_{t-k} + \varepsilon_t$$

2. Obtener los cuadrados del error de predicción y determinar su varianza descriptiva:

$$\hat{\sigma}^2 = \sum_{i=1}^N \hat{\varepsilon}_i^2 / N$$

3. Calcular la autocorrelaciones simples con la siguiente expresión:

$$\rho_i = \frac{\sum_{t=i+1}^N (\hat{\varepsilon}_t^2 - \hat{\sigma}^2)(\hat{\varepsilon}_{t-1}^2 - \hat{\sigma}^2)}{\sum_{t=1}^N (\hat{\varepsilon}_t^2 - \hat{\sigma}^2)}$$

4. Comprobar si son significativamente diferentes de cero a partir de la expresión de Ljung-Box con "n" grados de libertad. Rechazar la hipótesis nula de incorrelación en el valor de los residuos al cuadrado es equivalente a aceptar la existencia de un modelo ARCH o GARCH.

$$L-Q = N(N+2) \sum \rho_i / (N-i)$$



Esta expresión se distribuye como una  $\chi^2$  con "n" grados de libertad.

De cara a exponer un sistema de contraste de estructura estricto, Bollerslev propone aplicar el contraste del Multiplicador de Lagrange bajo la hipótesis nula de no existencia del proceso GARCH o, lo que es lo mismo, que todos los parámetros de la expresión de la varianza  $h_t$  son nulos, salvo la constante. Para ello, se construye el siguiente estadístico (particionando la matriz  $w = w_1 + w_2$  de forma que queden separadas constante y resto de parámetros). El sistema sería:

$$h_t = z_t' w = z_{t1}' w_1 + z_{t2}' w_2 \quad \text{Ec. 66}$$

Para la hipótesis nula  $H_0 : w_2 = 0$ , el contraste del multiplicador de Lagrange sería:

$$\xi_{LM}^* = \frac{1}{2} f_0' Z_0 (Z_0' Z_0)^{-1} Z_0' f_0 \quad \text{Ec. 67}$$

donde  $f_0$  es la vector del proceso al cuadrado normalizado (divido por la varianza) y  $Z_0$  es el vector de la varianza evaluada en la primera derivada, ambos para cada una de las distintas observaciones:

$$\begin{aligned} f_0 &= (\varepsilon_1^2 h_1^{-1} - 1, \dots, \varepsilon_T^2 h_T^{-1} - 1)' \\ Z_0 &= \left( h_1 \frac{\partial h_1}{\partial w}, \dots, h_T \frac{\partial h_T}{\partial w} \right) \end{aligned} \quad \text{Ec. 68}$$

Cuando la hipótesis nula es cierta, el valor de este contraste se distribuye como una  $\chi_r^2$ , siendo "r" el número de parámetros incluidos en  $w_2$ .

El empleo de este sistema ofrece resultados ligeramente diferentes a los que resultarían de la operación  $T \cdot R^2 \Rightarrow \chi_r^2$  de Breusch y Pagan (1978). Ello es debido a que la derivada parcial de la varianza respecto a los parámetros no se simplifica cuando la ecuación de la varianza condicional contiene términos desplazados de la propia varianza.

Finalmente, Bollerslev hace las siguientes consideraciones sobre el contraste expuesto:

- Bajo la hipótesis nula,  $Z_0'Z_0$  es una matriz singular si tanto p como q son mayores que cero, por lo que no se puede construir un test general para el proceso GARCH(p,q).
- En el caso de un ARCH(q) como hipótesis nula,  $Z_0'Z_0$  es singular para un GARCH( $r_1, q+r_2$ ) para  $r_1 > 0$  y  $r_2 > 0$ .
- En el ARCH(q) como hipótesis nula, el test LM para el GARCH(r,q) y el ARCH(q+r) coinciden.

#### 4.5. Predicción

En el modelo GARCH(p,q), los valores de predicción de la varianza condicional se pueden definir del siguiente modo:

Cuadro 6: Valores de Predicción de la Varianza Condicional

$h_{t+s} = \alpha_0 + \sum_{i=1}^n \alpha_i \varepsilon_{t+s-i}^2 + \beta_1 h_{t+s-1} + \sum_{i=2}^m \alpha_i \varepsilon_{t+s-i}^2 + \beta_i h_{t+s-i}$ $n : \min\{m, s-1\}; \quad m : \max\{p, q\} \quad s > m$		
	<b>Marginal</b>	<b>Condicional</b>
<b>Esperanza</b>	$E(\varepsilon_{t+s}^2) = \sigma_\varepsilon^2 = \alpha_0 \left[ 1 - \sum (\alpha_i + \beta_i) \right]^{-1}$	$E_t(h_{t+s}) = \alpha_0 + \sum_{i=1}^n (\alpha_i + \beta_i) E_t(h_{t+s-i}) + \sum_{i=s}^m \alpha_i \varepsilon_{t+s-i}^2 + \beta_i h_{t+s-i}$

Para poder comparar las varianzas marginal y condicional, si se sustituye el término independiente de la segunda por  $\sigma_\varepsilon^2 (1 - \sum \alpha_i + \beta_i) = \alpha_0$  se obtiene la siguiente expresión:

$$E_t(h_{t+s}) = \sigma_\varepsilon^2 + \left( \sum_{i=1}^n (\alpha_i + \beta_i) E_t(h_{t+s-i}) - \sigma_\varepsilon^2 \right) + \sum_{i=s}^m (\alpha_i \varepsilon_{t+s-i}^2 + \beta_i h_{t+s-i}) \quad \text{Ec. 69}$$

De la ecuación Ec. 69 se pueden concluir que a medida que  $s \rightarrow \infty$ , la  $E(h_{t+s}) \rightarrow \sigma_\varepsilon^2$ .

Baillie y Bollerslev (1992) derivan la siguiente expresión del error cuadrático medio:

$$E_T(v_{T,s}^2) = (k_2 - 1)\alpha^2 \sum_{i=1}^{s-1} (\alpha + \beta)^{2(1-i)} E_T(\sigma_{t-s-i}^4)$$

En esta expresión,  $v_{T,s} = \sigma_{T+s}^2 - E_T(\sigma_{T+s}^2)$  y  $k_2$  es el valor del coeficiente de curtosis (el cumulante de orden 2 de la densidad condicional de  $\varepsilon_t$ , que, si es una normal, sería tres)

## **5. VARIANTES SOBRE LOS MODELOS AUTORREGRESIVOS HETEROCEDÁSTICOS CONDICIONALES**

## 5.1. Modelo IGARCH y Modelos de Memoria Larga (Component ARCH)

### Introducción

Las siglas IGARCH se corresponden con *Integrated Generalized Autoregressive Conditional Heteroskedastic*. El propósito de esta variante de los modelos ARCH(q) y GARCH(p,q) es estimar la varianza en el caso en que ésta es integrada en varianza. Como resultado, se llega a una estimación máximo-verosímil a partir de una t-student de la cual no se conocen los grados de libertad, hecho por el cual éstos también serán estimados.

El desarrollo seminal de este tipo de modelos se debe a Engle y Bollerslev (1986), autores que lo desfinen del siguiente modo:

Dado un proceso GARCH(p,q) determinado por su dos primeros momentos:

$$\begin{aligned} E_{t-1}(\varepsilon_t) &= 0 \\ E_{t-1}(\varepsilon_t^2) &= h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad \text{Ec. 70} \\ \alpha_0 &\geq 0, \alpha_i > 0 \text{ y } \beta_i > 0 \end{aligned}$$

En el que la suma de todos los parámetros es igual a la unidad (en forma polinomial):

$$1 - \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} = 0 \quad \text{Ec. 71}$$

Se dice que el proceso tiene  $d > 0$  raíces unitarias y  $\max(p,q) - d$  raíces fuera del círculo unitario. El proceso entonces será:

- a) Si  $\alpha_0 > 0$ , integrado en varianza de orden "d".
- b) Si  $\alpha_0 = 0$ , integrado en varianza con tendencia.

Los modelos IGARCH son un caso específico dentro de la familia de los "modelos con varianza persistente" en los que la información actual (en el momento "t") es importante para realizar predicciones óptimas a cualquier horizonte temporal (hecho consustancial a los modelos de heterocedasticidad condicional autorregresiva).

Para los valores de futuro del modelo IGARCH(1,1) podríamos escribir la esperanza condicional en el horizonte "t+1" conocidos los valores hasta "t" como sigue:

$$E_t(h_{t+1}) = E_t(\alpha_0 + \alpha_1 \varepsilon_t^2 + (1 - \alpha_1)h_t) = \alpha_0 + (1 - \alpha_1)h_t \quad \text{Ec. 72}$$

y, a partir de la ley de las expectativas iteradas, generalizar la expresión de la Ec. 72 para el momento de predicción "t+s" como:

$$E_t(h_{t+s}) = s\alpha_0 + h_{t+1} \quad \text{Ec. 73}$$

El modelo así obtenido es muy similar al "paseo aleatorio" con deriva tradicional. A medida que se aumenta el horizonte de predicción, es evidente que el peso de  $h_{t+1}$  en la expresión anterior es cada vez más pequeño, pero, aún así, nunca deja de aparecer, lo que permite hablar de "varianza persistente" en la terminología de Engle y Bollerslev.

Siguiendo la escritura del modelo GARCH(1,1) que propusiera Pantula (1984) y que ya se comentó en la sección 1.2, definiendo  $v_t = \varepsilon_t^2 - h_t$ , se puede escribir el mismo modelo para la primera diferencia con el fin de contrastar la existencia de una raíz unitaria:

$$\nabla \varepsilon_t^2 = -(1 - \alpha_1)(\varepsilon_{t-1}^2 - h_{t-1}) + (\varepsilon_{t-1}^2 - h_{t-1}) = -(1 - \alpha_1)v_{t-1} + v_t \quad \text{Ec. 74}$$

de tal modo que, si el verdadero modelo es  $h_{t+1} = \alpha_0 + \alpha_1 \varepsilon_t^2 + (1 - \alpha_1)h_t$ , las funciones de autocorrelación total y parcial para la primera diferencia del proceso  $\nabla \varepsilon_t^2$  no deben ser significativas. Para contrastar este hecho, Engle y Bollerslev proponen el del contraste de Dickey y Fuller (1979) en la determinación del parámetro  $(1 - \alpha_1)$

De cara a la estimación, dada la frecuente aparición de valores de la distribución del la varianza de la perturbación aleatoria con una cola superior a la de la normal (leptocúrticos), Engle y Bollerslev se plantearon en 1986 la posibilidad de que realmente no fuera la normal la verdadera distribución, hecho que, a su vez, haría replantearse el sistema de estimación.

Los autores comentados llegan a la conclusión de que el cuarto momento sigue una distribución más similar al de la t-student estandarizada con grados de libertad desconocidos, pero estimables. Esta conclusión se circunscribe a casos de muestras relativamente pequeñas (menos de 25 datos), asumiéndose que, para muestras más amplias, la función es indistinguible de la normal.

En definitiva, los sistemas de estimación ya comentados en los apartados referidos a los modelos ARCH y GARCH siguen siendo de perfecta aplicación para el caso en que contemos con muestras suficientes.

### **Modelos de memoria larga o ARCH de componentes**

La persistencia en volatilidad ha dado lugar a muy diversas interpretaciones y propuestas de solución. Recientemente, destaca la de Ding y Granger (1996), que da lugar a los procesos ARCH de Memoria Larga o ARCH de Componentes que, por su interés, resumimos.

Centrándonos en el proceso GARCH(1,1) podríamos definir teóricamente la función de autocorrelación que este proceso lleva implícito tal y como se hacía en su capítulo correspondiente y como sugieren Ding y Granger (1996). Recuérdese que Bollerslev (1986) derivaba un resultado similar al de la ecuaciones de Yule-Walker para un proceso ARIMA y el resultado para un GARCH (p,q) era:

$$\begin{aligned}\rho_k &= \sum_{i=1}^m \phi_i \rho_{k-i} & k \geq q+1 \\ \phi_i &= \alpha_i + \beta_i & \\ m &= \max(p, q)\end{aligned}\quad \text{Ec. 75}$$

Asumiendo estacionariedad en el proceso; es decir, para un GARCH (1,1) que  $\alpha + \beta < 1$ , la forma asintótica del proceso GARCH(1,1) se puede representar como:

$$h_t = \sigma^2(1 - \alpha - \beta) + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} \quad \text{Ec. 76}$$

donde  $\sigma^2$  es la varianza marginal del proceso. Reordenando, se puede obtener:

$$\varepsilon_t^2 - \sigma^2 = (\alpha + \beta)(\varepsilon_{t-1}^2 - \sigma^2) - \beta h_{t-1}(\varepsilon_{t-1}^2 - 1) + h_t(\varepsilon_t^2 - 1) \quad \text{Ec. 77}$$

Para determinar el primer valor de la función de autocovarianza se debe tomar la esperanza matemática  $E((\varepsilon_t - \sigma^2)(\varepsilon_{t-1} - \sigma^2)) = \gamma_1$ , lo que, aplicado al resultado anterior, sería lo mismo que multiplicar ambos lados por  $(\varepsilon_{t-1} - \sigma^2)$  y tomar esperanzas, es decir:

$$\gamma_1 = (\alpha + \beta)\gamma_0 - 2\beta E(h_{t-1}^2)$$

Como se expuso, la condición de existencia del momento de cuarto orden se da cuando  $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$ , y, como demuestra Bollerslev (1988):

$$\rho_1 = \alpha + \frac{\alpha^2\beta}{1 - 2\alpha\beta - \beta^2} \quad \text{Ec. 78}$$

Uniendo Ec. 75 con lo dicho más arriba (Ec. 78), tenemos que la función de autocorrelación sería:

$$\rho_k = \left( \alpha + \frac{\alpha^2\beta}{1 - 2\alpha\beta - \beta^2} \right) (\alpha + \beta)^{k-1} \quad \text{Ec. 79}$$



La función de autocorrelación muestral de diversas series, sobre todo financieras, es decreciente rápidamente al comienzo hasta llegado un momento en el que dicho decrecimiento es muy lento, pero siempre manteniendo valores de autocorrelación claramente significativos hasta momentos del tiempo bastante lejanos. Este comportamiento es distinto al que se deduce de una función logarítmica exponencial regida por una curva logística, tal y como lo es la función de autocorrelación teórica que se define para un GARCH(1,1) según la Ec. 79 que acabamos de deducir. Con frecuencia, los procesos GARCH (1,1) presentan unos parámetros que, sumados, dan valores muy próximos a uno; es decir, la memoria del proceso es larga y, sin embargo, la función de autocorrelación calculable para el proceso se agota rápidamente.

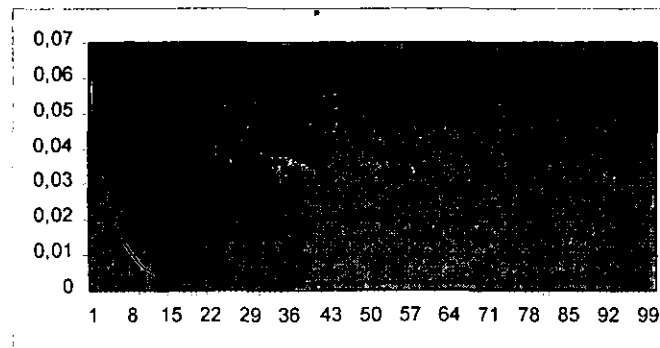
Este hallazgo investigador da lugar a la generación de un nuevo enfoque reciente en la familia ARCH que vendría a conocerse como ARCH de componentes o ARCH de Memoria Larga.

Escribiendo un proceso GARCH (1,1) como sucesión infinita de los términos del error, tendríamos:

$$h_t = w + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} = \frac{w}{1-\beta} + \alpha \sum_{k=1}^{\infty} \beta^{k-1} \varepsilon_{t-k}^2 \quad \text{Ec. 80}$$

En dicha expresión, la forma de logística del sumando no constante determina esa progresión rapidísima hacia la nulidad en el valor de significación de retardos muy próximos por efecto de contar con un valor de beta menor que uno.

#### Representación de la varianza condicional de un GARCH (1,1)



Parámetros, alfa = 0,2 y beta = 0,75

En el ejemplo, y a pesar de que la suma de los parámetros es prácticamente uno, la memoria del proceso GARCH (1,1) se hace casi nula antes de las diez primeras observaciones.

La propuesta del modelo ARCH de componentes se refiere a modelizar de diferente modo los efectos de corto plazo (imbuidos en un rápido decrecimiento que la forma exponencial del GARCH (1,1) recoge con relativo acierto) y otros de largo plazo, para los que se define una nueva función de comportamiento. Con este nuevo planteamiento se pretende dar solución no sólo a los distintos ritmos de aceleración presentes en las evolución de la varianza, sino también a la "persistencia" o larga memoria sobre su pasado que presenta. La varianza condicional del proceso quedaría así definida en dos ecuaciones diferentes cuyos valores estarían debidamente ponderados del siguiente modo<sup>8</sup>:

$$\begin{aligned}
 h_t &= wh_{1t} + (1-w)h_{2t} \\
 h_{1t} &= \alpha_1 \varepsilon_{t-1}^2 + (1-\alpha_1)h_{1t-1} \\
 h_{2t} &= h_t(1-\alpha_2-\beta_2) + \alpha_2 \varepsilon_{t-1}^2 + \beta_2 h_{2t-1}
 \end{aligned}
 \quad \text{Ec. 81}$$

En la formulación anterior, donde "w" es un factor de ponderación de cada una de las componentes, es inmediato ver que  $h_{1t}$  responde a un IGARCH (1,1) y  $h_{2t}$  responde a un

<sup>8</sup> La forma de escribir el término constante en la ecuación  $h_{2t}$  se corresponde con su escritura asintótica, en la que la varianza marginal coincide con la condicional (en ambos casos, toda la historia es conocida) y, por eso, al valor de la varianza marginal  $\sigma^2$  se le resta la proporción que se suma después en cada una de las componentes.

GARCH (1,1). Ambas expresiones se pueden escribir como suma infinita de los términos del error del siguiente modo:

$$\begin{aligned} h_{1t} &= \alpha_1 \sum_{k=1}^{\infty} (1 - \alpha_1)^{k-1} \varepsilon_{t-k}^2 \\ h_{2t} &= \sigma^2 \frac{1 - \alpha_2 - \beta_2}{1 - \beta_2} + \alpha_2 \sum_{k=1}^{\infty} \beta_2^{k-1} \varepsilon_{t-k}^2 \\ h_t &= \sigma^2 (1 - w) \frac{1 - \alpha_2 - \beta_2}{1 - \beta_2} + \sum [w \alpha_1 (1 - \alpha_1)^{k-1} + (1 - w) \alpha_2 \beta_2^{k-1}] \varepsilon_{t-k}^2 \end{aligned} \quad \text{Ec. 82}$$

Sustituyendo adecuadamente las ecuaciones que forman este ARCH por componentes o ARCH de memoria larga (como lo bautizaran sus creadores, - *long memory ARCH* -), estaríamos ante un GARCH (2,2) con parámetros restringidos.

$$\begin{aligned} h_t &= \sigma^2 (1 - w) \alpha_1 (1 - \alpha_2 - \beta_2) + [w \alpha_1 + (1 - w) \alpha_2] \varepsilon_{t-1}^2 \\ &\quad - [w \alpha_1 \beta_2 + (1 - w)(1 - \alpha_1) \alpha_2] \varepsilon_{t-2}^2 \\ &\quad + (1 - \alpha_1 - \beta_2) h_{t-1} - (1 - \alpha_1) \beta_2 h_{t-2} \end{aligned} \quad \text{Ec. 83}$$

A pesar de que el signo de los parámetros de segundo retardo son negativos, la condición para que el valor estimado de la varianza sea positivo es fácilmente deducible a partir de la suma de los parámetros de la Ec. 83:

$$\begin{aligned} \sum &= w \alpha_1 + (1 - w) \alpha_2 - w \alpha_1 \beta_2 - (1 - w)(1 - \alpha_1) \alpha_2 + \\ &\quad 1 - \alpha_1 + \beta_2 - (1 - \alpha_1) \beta_2 = \\ &\quad 1 - (1 - w) \alpha_1 (1 - \alpha_2 - \beta_2) \end{aligned} \quad \text{Ec. 84}$$

En la ecuación anterior, la suma es positiva siempre que el ponderador y el término  $\alpha_1$  estén entre cero y uno y que la suma  $0 < \alpha_2 + \beta_2 < 1$ . En estas condiciones el proceso es covarianza estacionario en la medida en que:

$$E(h_t)E(h_{1t}) = E(h_{2t}) = E(\varepsilon_t^2) = \sigma^2 \quad \text{Ec. 85}$$

Dado que no necesariamente se recoge perfectamente el patrón de comportamiento de la función de autocorrelación tomando dos componentes en la formación de la varianza condicional, los autores del ARCH de Memoria Larga proponen un modelo generalizado para N componentes cuya forma funcional sería la siguiente:

$$\begin{aligned} \varepsilon_t &= h_t \eta_t; & \eta_t &\rightarrow D(0,1) \\ h_t &= \sum_{i=1}^N w_i h_{it}; & \sum_{i=1}^N w_i &= 1 \\ h_{it} &= \sigma^2(1 - \alpha_i - \beta_i) + \alpha_i \varepsilon_{t-1}^2 + \beta_i h_{it-1} & i = 1 \dots N \end{aligned} \quad \text{Ec. 86}$$

Con esta formulación se dota a la varianza de "N" tendencias decrecientes diferentes posibles y estaríamos hablando, ahora, de un GARCH (N,N). Obviamente, el problema de un modelo tan general es que se desconoce el valor de "N" o del número de tendencias diferentes existente en la función de autocorrelación que se quiere replicar, y que la estimación puede ser muy compleja.

Para despejar estos problemas es necesario plantear una función de distribución de los parámetros. Los autores proponen la distribución Beta en el entorno [0,1], por cuanto esta es generalizable a otros casos sin gran dificultad. La función de densidad de Beta con parámetros p,q sería:

$$f(x) = \frac{x^{p-1} (1-x)^{q-1}}{B(p,q)} \quad \text{Ec. 87}$$

donde B(p,q) es la función Beta que se formula como:



$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx \quad \text{Ec. 88}$$

La forma funcional que se elija para los parámetros no es demasiado importante, tal y como demuestra Granger (1980), salvo para el caso en el que nos encontremos con valores extremos; es decir Beta igual a uno. De cara a la formulación, se asume que  $\alpha = (1 - \beta)\alpha^*$ , donde  $\alpha^*$  es un valor comprendido entre cero y uno (de media positiva). Es necesario también que  $0 \leq \alpha + \beta \leq 1$ . Bajo estas condiciones, la formulación para el proceso de N componentes (Ec. 86) se puede describir asintóticamente despejando los valores del residuo y obteniendo:

$$h_{it} = \sigma^2 \frac{1 - \alpha_i - \beta_i}{1 - \beta_i} + \frac{\alpha_i}{1 - \beta_i L} \varepsilon_{i,t-1}^2 \quad \text{Ec. 89}$$

Donde L es el operador retardo. Sustituyendo en el sumatorio de las N componentes, obtendríamos:

$$h_t = \sum_{i=1}^N w_i \left[ \sigma^2 \frac{1 - \alpha_i - \beta_i}{1 - \beta_i} + \frac{\alpha_i}{1 - \beta_i L} \varepsilon_{i,t-1}^2 \right] \quad \text{Ec. 90}$$

A partir de la definición de la función de distribución de los parámetros que se ha supuesto en Ec. 87, cuando N tiende a infinito, tenemos:

$$\begin{aligned}
h_t &= \sigma^2 \int \frac{1-\alpha-\beta}{1-\beta} dF(\alpha, \beta) + \int \frac{\alpha}{1-\beta L} \varepsilon_{t-1}^2 dF(\alpha, \beta) \\
&= \sigma^2 \int (1-\alpha^*) dF(\alpha^*) + \sum_{k=1}^{\infty} \varepsilon_{t-k}^2 \int \alpha^* dF(1-\beta) \beta^{k-1} dF(\alpha^*, \beta) \\
&= \sigma^2 (1-\mu) + \sum_{k=1}^{\infty} \varepsilon_{t-k}^2 \int \alpha^* dF(\alpha^*) \int (1-\beta) \beta^{k-1} dF(\beta) = \\
&= \sigma^2 (1-\mu) + \mu \sum_{k=1}^{\infty} \frac{B(p+k-1, q+1)}{B(p, q)} \varepsilon_{t-k}^2
\end{aligned}$$

Ec. 91

El proceso que queda definido según la ecuación anterior es claramente de "memoria larga" a juzgar por su parámetro sobre las innovaciones previas, que sería aproximadamente igual a:

$$\begin{aligned}
a_k &= \frac{B(p+q-1, q+1)}{B(p, q)} = \frac{q\Gamma(p+q)\Gamma(k-1)}{\Gamma(p)\Gamma(p+q+k)} \\
&\approx \frac{q\Gamma(p+q)}{\Gamma(q)} k^{-1-q}
\end{aligned}$$

Ec. 92

Donde, salvo para el valor extremo de  $p=0^9$ , el patrón de comportamiento de decrecimiento del parámetro del modelo (que hemos llamado  $a_k$ ) viene únicamente determinado por la evolución de "q".

Con el modelo de N componentes se logra que el decrecimiento del parámetro de referencia se vaya reduciendo según una función hiperbólica, en vez de exponencialmente como lo hace en el caso de un GARCH (1,1). En esta misma línea se centra el modelo FIGARCH o GARCH integrado fraccional que propusieron Granger y Ding (1995).

<sup>9</sup> En este caso ( $p=0$ ), la función anterior se convierte en un ARCH (1) y no hay proceso de decrecimiento paulatino o de memoria larga, sino que el proceso se agota tempranamente como ya se ha comentado previamente.

## **5.2. Modelo E-GARCH**

### **Introducción**

Los modelos GARCH exponenciales nacen a partir de la publicación de Daniel Nelson (1991) sobre heterocedasticidad condicional en los modelos de rentabilidad de activos. Dicho autor propone una nueva forma funcional a partir de la observación de las deficiencias de los ya ampliamente actualizados de los modelos tipo ARCH. En su artículo critica tres elementos de los procesos GARCH:

- las restricciones de no negatividad de los parámetros son difíciles de lograr en muchas ocasiones.
- Los modelos GARCH no permiten estimar convenientemente el efecto de apalancamiento financiero que aparece en la realidad.
- Los modelos IGARCH son difíciles de llevar a la práctica, siendo confuso el término de persistencia en varianza condicional acuñado por Engle y Bollerslev.

En primer lugar, Nelson cita la importancia de considerar diferentes efectos en la estimación de la varianza condicional según la innovación en el período anterior sea negativa o positiva. Es decir, es necesario hablar de efectos asimétricos en función del signo de la perturbación en el período anterior, respondiendo a lo empíricamente observado, por ejemplo, en los mercados financieros: se produce un aumento de la volatilidad mayor cuando la "innovación" del período previo fue negativa que cuando la "innovación" en el período anterior fue positiva.

En segundo lugar, Nelson critica el efecto expansivo de los modelos ARCH tradicionales. El hecho de que los parámetros deban ser necesariamente positivos en las formulaciones de ARCH y GARCH, produce un efecto explosivo y siempre creciente en la estimación a futuro. La realización de un valor de la varianza condicional positiva en un período " $t$ " ( $\varepsilon_t^2$ ) produce una estimación siempre positiva para la estimación de  $h_{t+m}$  para cualquier  $m \geq 1$ , eliminando la posibilidad de un comportamiento aleatorio oscilatorio en el futuro.

En tercer lugar, Nelson agrega que la necesidad de contar con parámetros positivos produce dificultades para la estimación, como ya ponían de manifiesto Engle, Lilien y Robins (1987), imponiendo al sistema ARCH una estructura de  $\alpha_i$  parámetros ponderados de forma decreciente de cara a realizar la estimación.

Por último, los modelos ARCH tradicionales encuentran dificultades en la medición de la persistencia de un "shock" pasado en los valores de la varianza condicional. En diferentes estudios citados por el mismo Nelson, se pone de manifiesto que si el impacto en la volatilidad de un determinado "shock" persiste indefinidamente, se transformaría completamente la estructura temporal de la rentabilidad, produciendo un fuerte impacto sobre las inversiones en bienes de larga duración. (sobre este particular, destaca el estudio de Poterba y Summers (1986)).

#### **Especificación del modelo GARCH exponencial (EGARCH)**

Fruto de las críticas antes comentadas a las deficiencias de los modelos tipo ARCH tradicionales para recoger la realidad empírica suministrada por las variables financieras, Nelson propone una nueva especificación más general sobre la varianza condicional heterocedástica que da nombre a los GARCH exponenciales o EGARCH (p,q).

Al igual que los modelos ARCH tradicionales aseguran la estimación de valores positivos en la varianza condicional con una estructura lineal de dependencia que incorpora parámetros positivos sobre variables explicativas aleatorias positivas (tanto los residuos al cuadrado como la propia varianza condicional desplazada lo son), Nelson propone la estimación del modelo en logaritmos, con lo que también existe probabilidad uno de obtener valores de la varianza positivos, aún cuando algún o algunos parámetros no sean positivos:

$$y_t = \varepsilon_t h_t^{1/2}$$

$$\log h_t = \alpha_0 + \sum_{k=1}^q \beta_k g(\varepsilon_{t-k}) \quad \beta_1 \equiv 1$$

$$g(\varepsilon_t) = \theta \varepsilon_t + \gamma [|\varepsilon_t| - E(|\varepsilon_t|)]$$



Con esta formulación quedan corregidas las deficiencias que, sobre el modelo tipo ARCH tradicional, señalaba el autor y que se comentaban en el epígrafe anterior:

- Por un lado, ya ha quedado patente la cobertura de la necesidad estadística de obtener valores de la varianza estimada siempre positivos al realizar una formulación en logaritmos, hecho que ya se había producido en los comentarios de Pantula (1986) y Geweke (1986).
- La respuesta asimétrica al diferente signo de la innovación previa en el modelo viene perfectamente recogida por la especificación dada a la función  $g(\cdot)$  anterior. Para valores negativos de  $\varepsilon_t$ , la función  $g(\varepsilon_t)$  es lineal y con tendencia  $(\theta + \gamma)$  y, para valores positivos, es  $(\theta - \gamma)$ . El segundo sumando de la definición de la función  $g(\cdot)$  es el determinante de la asimetría ya que marca la diferencia entre el valor realizado de la innovación y su valor esperado. Esta diferencia obviamente puede ser positiva o negativa, incorporando mayor o menor volatilidad (mayor o menor varianza condicional) en función de su signo.
- la objeción hecha a los modelos ARCH tradicionales de no permitir comportamientos a futuro oscilantes queda salvada en esta especificación logarítmica, donde los parámetros no tienen porqué ser positivos en todos los casos.
- Respecto a la persistencia de los "shocks", la estructura lineal del modelo planteado permite contrastar fácilmente la estacionariedad y la ergodicidad. El  $\ln(h_t)$  será estrictamente ergódico y estacionario en la medida en que el efecto o "shock" de innovaciones pasadas se elimine lo suficientemente rápido. La condición de estacionariedad de este proceso es la misma que para un ARMA de orden infinito (la especificación dada no es más que un modelo de medias móviles de orden infinito), con lo cual basta con comprobar que los valores de los parámetros obtenidos sean, en suma, menor que uno.

De cara a mantener la misma signatura que en el resto de los modelos comentados, lo dicho anteriormente, escribiendo los valores estandarizados de las perturbaciones aleatorias, se puede describir del siguiente modo para lo que sería un EGARCH(1,1):

$$y_t = \varepsilon_t h_t^{1/2}$$

$$\log h_t = \alpha_0 + \beta_1 \log(h_{t-1}) + \alpha_1 \left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}$$

donde podremos distinguir claramente varios componentes:

- el de la varianza heterocedástica:  $\beta_1 \log(h_{t-1})$
- el del valor autorregresivo de la perturbación:  $\gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}$
- el efecto asimétrico:  $\left| \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \right| i$ , donde  $i = \begin{cases} 1 & \text{si } \varepsilon_{t-1} < 0 \\ 0 & \text{si } \varepsilon_{t-1} > 0 \end{cases}$

Con ello se pone de manifiesto que la formulación de Nelson parte el valor del parámetro que acompaña a la perturbación aleatoria en aquel debido al signo y aquel debido a su propia naturaleza.

La segunda gran aportación del documento seminal de los modelos EGARCH hace referencia a la función de distribución de la perturbación aleatoria. Con frecuencia, las series a las que se aplica el modelo tipo ARCH muestran una distribución con mayor apuntamiento que el de una normal y con colas "más gruesas". Es por ello que se plantea aquí una forma de la función de densidad del proceso que admita, como caso especial, una distribución normal; pero que, al tiempo, permita otros tipos de funciones.

Escrito el modelo EGARCH como un proceso ARMA obtendríamos la siguiente expresión:

$$\ln(h_t) = \alpha_0 + \frac{(1 + \psi L + \dots + \psi_q L^q)}{(1 - \Delta_1 L - \dots - \Delta_p L^p)} g(\varepsilon_{t-1})$$

donde se asume que tanto  $1 - \sum_{i=1, \dots, p} \Delta_i y^i$  como  $1 + \sum_{i=1, \dots, q} \psi_i y^i$  no tienen raíces unitarias.

Con ello, la varianza estimada que se sigue de la expresión anterior ( $\exp(-\alpha_0) \cdot h_t$ ) y

$\exp(-\alpha_i/2) \cdot \varepsilon_i$  serán estrictamente estacionarios y ergódicos siempre que todas sus raíces unitarias estén fuera del círculo unitario.

La estacionariedad de los momentos condicionales no implica necesariamente lo mismo para los momentos marginales. Por ejemplo, una t-student con un número de grados de libertad determinado, presenta media y varianzas infinitas para los valores de  $\exp(-\alpha_i) \cdot h_i$  y de  $\exp(-\alpha_i/2) \cdot \varepsilon_i$ . Por esta razón, Nelson propone emplear una función de distribución del proceso del tipo de la definida por Harvey (1981) - GED: *Generalized Error Distribution* -, que para nuestro caso podría escribirse del siguiente modo:

$$f(\varepsilon_i) = \frac{v \cdot \exp\left[-\frac{1}{2} \left|\frac{\varepsilon_i}{\lambda}\right|^v\right]}{\lambda 2^{(1+1/v)} \Gamma(1/v)} \quad -\infty < \varepsilon < \infty; \quad -\infty < v < \infty$$

$$\lambda = \left[2^{(-2/v)} \Gamma(1/v) / \Gamma(3/v)\right]^{1/2}$$

donde  $\Gamma(\cdot)$  es la función gamma y  $v$  es la densidad de la cola del parámetro. Cuando  $v=2$ ,  $\varepsilon_i$  se distribuye según una normal. Si  $v < 2$ , la distribución es entonces más gruesa que la de la normal (por ejemplo, si  $v=1$ , la distribución es una exponencial doble) y, para  $v > 2$ , la distribución de  $\varepsilon_i$  es más delgada que la de la normal (por ejemplo, si  $v = \infty$ ,  $\varepsilon_i$  presenta una distribución uniforme en el intervalo  $(-3^{1/2}, 3^{1/2})$ ).

### Contraste de especificación

La función de distribución que se ha supuesto exige los siguientes cumplimientos para comprobar las condiciones de ortogonalidad:

- La esperanza de  $\varepsilon_i$  debe ser nula
- La varianza marginal debe ser igual a uno
- La esperanza de la función  $g(\varepsilon_i)$  para también debe ser nula:  $E(g(\varepsilon_i)) = 0$
- Por simetría de la función GED, la  $E(\varepsilon_i | \varepsilon_i) = 0$

Además, en el modelo supuesto, es necesario que los valores de  $\varepsilon_t^2 - 1$  y  $\varepsilon_t$  estén serialmente incorrelacionados. Nelson propone la cumplimentación de la siguiente tabla, empleando como valores estimados las medias muestrales de cada uno de los valores que aparecen y la comprobación de su significatividad estadística en función de una t-student o de una chi-cuadrado.

Condiciones de ortogonalidad			
1	$E(\varepsilon_t) = 0$	8	$E[(\varepsilon_t^2 - 1)(\varepsilon_{t-4}^2 - 1)] = 0$
2	$E(\varepsilon_t^2 - 1) = 0$	9	$E[(\varepsilon_t^2 - 1)(\varepsilon_{t-5}^2 - 1)] = 0$
3	$E( \varepsilon_t    -\lambda 2^{1/\nu} \Gamma(2/\nu) \Gamma(1/\nu) = 0$	10	$E[\varepsilon_t \varepsilon_{t-1}] = 0$
4	$E(\varepsilon_t ( \varepsilon_t  - E( \varepsilon ))) = 0$	11	$E[\varepsilon_t \varepsilon_{t-2}] = 0$
5	$E[(\varepsilon_t^2 - 1)(\varepsilon_{t-1}^2 - 1)] = 0$	12	$E[\varepsilon_t \varepsilon_{t-3}] = 0$
6	$E[(\varepsilon_t^2 - 1)(\varepsilon_{t-2}^2 - 1)] = 0$	13	$E[\varepsilon_t \varepsilon_{t-4}] = 0$
7	$E[(\varepsilon_t^2 - 1)(\varepsilon_{t-3}^2 - 1)] = 0$	14	$E[\varepsilon_t \varepsilon_{t-5}] = 0$

Con las primeras nueve condiciones se pretende contrastar la correcta especificación del modelo y, con las cinco últimas, la ausencia de correlación serial en la variable que se pretende explicar. Para su validación estadística (para determinar si son o no significativamente distintas de cero, valor requerido) se emplean sus funciones de distribución; es decir, los valores tabulados de la  $\chi^2$  con el número de grados de libertad que corresponda a cada criterio (al número de condición en la tabla anterior). Para estimar las condiciones es frecuente utilizar la media muestral.

### **5.3. Modelos GJR - ARCH**

#### **Introducción**

Glosten, Jagannthan y Runkle (1993) propusieron un nuevo modelo tipo ARCH aplicable a ciertas constataciones sobre los mercados financieros y, en particular, sobre los mercados bursátiles. Su aportación, que daría nombre al modelo conocido como GJR-ARCH se centra en los siguientes aspectos:

Bajo el supuesto de que los precios de los activos se forman a partir de una ley de expectativas centradas en la volatilidad que sobre ellos se ha observado en el pasado (CAPM), el económetra se encuentra con fuertes restricciones de información a la hora de realizar sus predicciones sobre el valor futuro de dicha volatilidad. Siguiendo un esquema tipo ARCH (considerar que la volatilidad se autoalimenta de su propio pasado) daría lugar a dos tipos de errores: uno el del propio supuesto simplificador de la realidad, en la que seguro intervienen más variables explicativas; y, otro, el del sistema de estimación posible. Como sistema paliativo de estos problemas, los autores referidos diseñan un modelo de estimación en el que incorporan tres componentes nuevos en los habituales ARCH:

- Las variables financieras difieren en la forma de afrontar (volatilidad) "las buenas y las malas noticias": con frecuencia, los movimientos de caída son más bruscos y largos que los de alza. Una de las aportaciones del modelo GJR-ARCH será la consideración particular de este carácter asimétrico de las innovaciones.
- La volatilidad de los mercados es tanto menor cuanto mayor sea la prima de riesgo cubierta por estos. Con ello se quiere incidir en que un mercado con un alto nivel diferencial entre los activos con y sin riesgo es menos susceptible a cambios en el resto de los condicionantes generales de la economía, ya que, en cierto modo, estos ya estaban descontados en ese fuerte seguro que supone una alta prima de riesgo.
- Existen una serie de patrones estacionales observables que permiten aconsejar la inclusión de variables ficticias en los modelos de cartera de valores para incluir estos

"fenómenos conocidos". Los hábitos tanto personales como empresariales en el suministro de la información permiten observar altas volatilidades en los meses de enero y octubre en lo que a la bolsa se refiere.

- Por último, las expectativas de inflación modifican la volatilidad futura. Los autores del GJR-ARCH proponen los tipos de interés nominales a corto plazo como predictores de la inflación, entendiendo que, a partir de la ley de Fisher, éstos están incorporando en su formación el crecimiento previsible de los precios.

Dentro de la especificación econométrica que los autores del GJR-ARCH proponen, se incluirán entonces las variables y situaciones descritas como variables instrumentales en el proceso de estimación y predicción del verdadero valor de la varianza condicional. Posteriormente, a la matriz de estos instrumentos la llamaremos  $I_t$ , indicando claramente como se introducen estas variables en el modelo ARCH.

Para determinar el verdadero valor de una variable  $x_t$  se parte de que el valor esperado de esta. Dada la información disponible hasta  $t-1$  se propone la siguiente regla de formación o Proceso Generador de Datos:

$$E(z_t | F_{t-1}) = \alpha + \beta \text{var}(z_t | F_{t-1}) \quad \text{Ec. 93}$$

En este supuesto, obviamente se está incurriendo en una cierta inexactitud, a la que llamaremos  $\eta_t$ , a la hora de apreciar el verdadero valor realizado de  $x_t$ . Dicha inexactitud o error proviene de dos factores distintos:

1. la diferencia entre el valor esperado y el finalmente realizado:  $\varepsilon_t = z_t - E(z_t | F_{t-1})$ , a la que se le supone esperanza cero.
2. un segundo error proveniente de la mala observación de la información disponible hasta  $t-1$ . Es decir, es posible que, en vez de contar con el verdadero conjunto de información hasta  $t-1$  (que se ha llamado  $F_{t-1}$ ), trabajemos con uno distinto (que se llamará  $G_{t-1}$ ) y que la expresión de la varianza condicional que estamos utilizando como variable explicativa no sea la correcta, sino una aproximación de ésta. Podemos expresar esta segunda fuente de error en la variable explicativa del modelo como

$u_{t-1} = \beta(v_{t-1} - E(v_{t-1} | G_{t-1}))$ , donde por la Ec. 93,  $v_{t-1} = \text{var}(z_t | F_{t-1})$ . Con esta expresión, introducimos en el error aquella parte que supone la diferencia entre la verdadera explicativa del modelo teórico ( $v_{t-1}$ ) y aquella que nosotros estamos empleando  $E(v_{t-1} | G_{t-1})$ , sesgados por la información imprecisa sobre el pasado.

En definitiva, en el término de error  $\eta_t$  se agregan dos tipos de error: el del supuesto inicial (igualador de valor esperado y valor real) y el de emplear información deficiente sobre la varianza condicional de  $x_t$ .

$$\eta_t = [\beta(v_{t-1} - E(v_{t-1} | G_{t-1}))] + [x_t - E(z_t | F_{t-1})] = u_{t-1} + \varepsilon_t \quad \text{Ec. 94}$$

donde el valor real de  $z_t$  sería igual a:

$$z_t = \alpha + \beta E(x_t | G_{t-1}) + \eta_t \quad \text{Ec. 95}$$

La varianza condicional de  $z_t$  dada la información  $F_{t-1}$ , es decir  $v_{t-1}$ , se puede resumir en la siguiente expresión:

$$v_{t-1} = \text{var}(z_t | F_{t-1}) = \text{var}[(E(z_t | F_{t-1}) + \varepsilon_t) | F_{t-1}] = E(\varepsilon_t^2 | F_{t-1}) \quad \text{Ec. 96}$$

Con esta ecuación se puede escribir  $u_{t-1}$  como

$$u_{t-1} = \beta[E(\varepsilon_t^2 | F_{t-1}) - E(\varepsilon_t^2 | G_{t-1})] \quad \text{Ec. 97}$$

Seguidamente, la media condicional del error total del modelo  $\eta_t$ , se pueden determinar como:

$$E(\eta_t | G_{t-1}) = E(u_{t-1} | G_{t-1}) = E(u_{t-1} \varepsilon_{t-1}^2 | G_{t-1}) = 0 \quad \text{Ec. 98}$$

La varianza total del error de predicción del modelo se puede diferenciar en sus dos componentes:

$$E(\eta_t^2 | G_{t-1}) = E(u_{t-1}^2 | G_{t-1}) + E(\varepsilon_{t-1}^2 | G_{t-1}) \quad \text{Ec. 99}$$

error total
error medida
error modelo

### Método de estimación por variables instrumentales de Campbell (1987)

Campbell propone un modelo de variables instrumentales que sirva para contrastar el supuesto de partida de constancia arbitraria en el valor de la  $\text{var}(v_{t-1} | G_{t-1})$ , circunstancia que haría posible la estimación del modelo expresado anteriormente.

Si escribimos una regla de evolución de la esperanza de la varianza del siguiente modo:

$$E(v_{t-1} | G_{t-1}) = b_0 + b_1 l_{t-1} \quad \text{Ec. 100}$$

donde  $z$  es una representación de  $G$  y  $l_{t-1}$  es un vector columna y  $b_1$  un vector fila, podemos escribir la varianza del error total del modelo como:

$$E(\eta_t^2 | l_{t-1}) = \beta^2 \text{var}(v_{t-1} | G_{t-1}) + (b_0 + b_1 l_{t-1}) \quad \text{Ec. 101}$$

Si suponemos ahora constancia en el valor de  $\text{var}(v_{t-1} | G_{t-1})$ ,  $l_{t-1}$  se convierte en un vector de variables deterministas observables, con lo que se puede realizar la estimación de las dos ecuaciones siguientes:

$$\begin{aligned} x_t &= c_0 + c_1 l_{t-1} + \eta_t \\ \eta_t^2 &= d_0 + d_1 l_{t-1} + \zeta_t \end{aligned} \quad \text{Ec. 102}$$

El valor estimado de  $c_1$  equivale al de  $\beta b_1$  y el de  $d_1$  a al de  $b_1$ , por lo que de su cociente obtenemos el valor estimado del parámetro beta.



### Modelo GJR-ARCH: Modificaciones sobre el modelo GARCH-M de Bollerslev

Si expresamos la dependencia de una variable  $x_t$  de las realizaciones de la varianza de la innovación en el pasado y estas las definimos a su vez como un modelo GARCH (1,1), podríamos expresar el siguiente sistema de evolución:

$$\begin{aligned} z_t &= \alpha_0 + \alpha_1 h_{t-1} + \varepsilon_t \\ h_{t-1} &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-2} \end{aligned} \quad \text{Ec. 103}$$

donde  $E_{t-1}(\varepsilon_t) = 0$  y  $h_{t-1} = E_{t-1}(\varepsilon_{t-1}^2)$ .

En este modelo se supone implícitamente que la respuesta de la varianza condicional es la misma ante el mismo valor realizado de la innovación, sin realizar ninguna distinción por el signo de ésta. De cara a introducir el carácter asimétrico que la observación de la realidad parece indicar, el modelo GJR-ARCH propone la siguiente especificación alternativa de la varianza de la Ec. 103:

$$\begin{aligned} z_t &= \alpha_0 + \alpha_1 h_{t-1} + \varepsilon_t \\ h_{t-1} &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2 I_{t-1} + \beta_1 h_{t-2} \end{aligned} \quad \text{Ec. 104}$$

donde la variable ficticia  $I_{t-1}$  toma el valor cero cuando la innovación realizada en  $t-1$  es negativa y uno en el caso contrario.

De cara a incorporar más información sobre el pasado diferente a las de las propias realizaciones de la innovación del modelo GARCH-M, los autores del GJR-ARCH proponen la introducción de las variables explicativas que ya se comentaban anteriormente en la explicación de la varianza condicional.

Para recoger efectos estacionales en la especificación de la varianza condicional, la propuesta realizada por Glosten, Jagannathan y Runkle se centra en el principio de que la

verdadera variable explicativa de la varianza es la innovación desestacionalizada a la que se la denomina  $\eta_t$  y que guarda la siguiente relación con la innovación observada:

$$\varepsilon_t = (1 + \lambda_1 EST1_t + \lambda_2 EST2_t) \eta_t \quad \text{Ec. 105}$$

Siendo  $EST1_t$  y  $EST2_t$  las variables ficticias representativas de los meses en los que se supone una estacionalidad significativamente distinta en cuanto a la importancia de la volatilidad (en el estudio comentado, enero y octubre) y  $\lambda_1$  y  $\lambda_2$  los factores de escala de estacionalidad que se estimen.

El modelo propuesto, quedaría entonces del siguiente modo:

$$\begin{aligned} z_t &= a_0 + a_1 h_{t-1} + \varepsilon_t \\ h_{t-1} &= \alpha_0 + \alpha_1 \eta_{t-1}^2 + \alpha_2 \eta_{t-1}^2 I_{t-1} + \beta_1 h_{t-2} \end{aligned} \quad \text{Ec. 106}$$

En algunos estudios sobre la incorporación de la estacionalidad como el de Baillie y Bollerslev (1989) la forma en la que esta se introduce en el modelo afecta directamente a la especificación de la varianza. En la especificación descrita, el valor que se introduce en la ecuación de la varianza,  $\eta_t$ , está desestacionalizado.

La posibilidad de que los coeficientes lambda de la especificación presentada sean, sumados, un número negativo, deja lugar a la entrada de un resultado de la varianza condicional negativo (absurdo) en función de determinados valores de la innovación. Es por ello que se propone una especificación logarítmica adicional. También es constatable empíricamente que el efecto de una fuerte innovación (en valor absoluto) en el período anterior no produce un efecto proporcionalmente más grande que el de una débil. Para minimizar el efecto de estos grandes sobresaltos en la reacción del período posterior, se incluye en el modelo el término  $\eta_t / \sqrt{h_{t-2}}$ , tal y como propone Nelson (1991). La especificación resultante de incorporar estos hechos sería:

$$\begin{aligned} z_t &= a_0 + a_1 h_{t-1} + \varepsilon_t \\ \log(h_{t-1}) &= \alpha_0 + \alpha_1 \eta_{t-1}^2 / \sqrt{h_{t-2}} + \alpha_2 \eta_{t-1}^2 I_{t-1} / \sqrt{h_{t-2}} + \beta_1 \log(h_{t-2}) \end{aligned} \quad \text{Ec. 107}$$

En definitiva, los diferentes modelos y su aportación diferencial van incorporando paulatinamente algún tipo de información en la explicación de la varianza condicional y se podrían resumir en el siguiente cuadro.

**Modelos propuestos en Glosten y otros (1992)**

<b>Modelos</b>	<b>Especificación</b>	<b>Comentario</b>
<b>Modelo I</b> <b>GARCH-M</b>	$z_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $h_{t-1} = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-2}$	Modelo tradicional de Bollerslev
<b>Modelo II</b> <b>GARCH-M</b> <b>Ampliado</b>	$z_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $h_{t-1} = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-2} + \beta_2 r_f$	Introducción de más variables explicativas en la varianza condicional (por ejemplo, interés sin riesgo)
<b>Modelo III</b> <b>Shocks</b> <b>Asimétricos</b>	$z_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $h_{t-1} = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2 I_{t-1} + \beta_1 h_{t-2}$	Diferenciación del impacto de innovaciones pasadas negativas o positivas
<b>Modelo IV</b> <b>Factores</b> <b>Estacionales</b>	$z_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $h_{t-1} = \alpha_0 + \alpha_1 \eta_{t-1}^2 + \alpha_2 \eta_{t-1}^2 I_{t-1} + \beta_1 h_{t-2}$ $\varepsilon_t = (1 + \lambda_1 EST1_t + \lambda_2 EST2_t) \eta_t$	Incorporación de los efectos estacionales típicos de los mercados bursátiles
<b>Modelo V</b> <b>Estructura</b> <b>E-GARCH</b>	$z_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $\log(h_{t-1}) = \alpha_0 + \alpha_1 \eta_{t-1}^2 / \sqrt{h_{t-2}} + \alpha_2 \eta_{t-1}^2 I_{t-1} / \sqrt{h_{t-2}} + \beta_1 \log(h_{t-2})$	Forma logarítmica (forzar valores positivos) y término EGARCH (romper la proporcionalidad en el impacto de innovaciones anormalmente altas)

## Estimación y contraste

De cara a la estimación de la especificación propuesta, el punto de partida es el común a todos los modelos tipos ARCH y éste es el supuesto de la distribución normal de la varianza condicional. Así, los autores proponen una estimación consistente de los parámetros a partir del logaritmo de la función de máxima verosimilitud. Cuando la asunción sobre la distribución no sea correcta (hecho que ha de comprobarse en la fase de contraste), la estimación es posible a partir de métodos como el de cuasi-máxima verosimilitud, obteniéndose resultados consistentes ya que la distribución será asintóticamente normal (demostrado en Bollerslev y Wooldrige (1992) o Glosten y otros (1988)).

Para el contraste de las especificaciones realizadas, los autores del GJR-ARCH proponen tres validaciones diferentes:

1. Se contrasta la normalidad de los residuos a partir de los coeficientes de asimetría y curtosis. En principio, los modelos aplicados debieran corregir el carácter leptocúrtico que suelen manifestar las series financieras.
2. Se contrasta si los residuos al cuadrado estandarizados son independientes e idénticamente distribuidos. Para ello se emplea la curva de "impacto de la noticias" expresada por Engle y Ng (1993). Con ello se contrasta al tiempo la forma en la que afectan diferencialmente innovaciones pasadas negativas o positivas y en qué medida afectan al valor futuro la cuantía absoluta de éstas.
3. Para contrastar la persistencia en varianza o el potencial hecho de que la varianza condicional presente una raíz unitaria, se realiza una regresión de  $h_t$  en función de una constante y de  $h_{t-1}$  con el fin de comprobar la significatividad del coeficiente y si este es superior a uno. Este contraste no riguroso permite descartar unos modelos frente a otros (de entre los propuestos en la sección anterior).

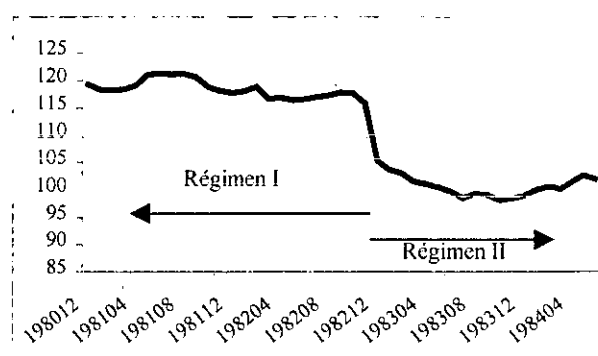
## 5.4. Modelos de cadenas de Markov y SW-ARCH (Switching-ARCH)

### Cambios de régimen con cadenas de markov

#### Introducción a las cadenas de Markov

Con gran frecuencia, las series económicas muestran estructuras diferentes separadas claramente en un momento del tiempo. Por ejemplo, la que aparece en la siguiente ilustración para el índice del tipo de cambio efectivo nominal de la peseta frente al Euro en los primeros cinco años de los ochenta, muestra dos estructuras diferentes: antes y a partir de la devaluación Boyer de diciembre de 1982.

En el denominado régimen I, la media del índice rondaría el 120 y, a partir del comienzo del régimen II, el 100, con un salto importante en el período de transición de un estado o régimen a otro (diciembre de 1982). Diversos acontecimientos como guerras, catástrofes naturales, creación de pactos colusorios de mantenimiento de precios, pánicos financieros, cambios dramáticos en política económica, simples cambios metodológicos, etc. son susceptibles de dar lugar a "cambios de régimen" tales como el que se puede ver en la ilustración.



Estadísticamente, está claro que sería más apropiado tener en cuenta dos medias diferentes, una para el régimen uno y otra para el régimen dos, para realizar una aproximación más verídica a los datos observables en el entorno de cada una de ellas. A cada uno de estos regímenes, en el marco de las denominadas Cadenas de Markov se les llama "estados" y su ocurrencia viene fijada por un comportamiento aleatorio cuya representación, en términos de probabilidad, se puede escribir con la siguiente matriz:

$$\begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} \quad \text{Ec. 108}$$

donde, cada elemento  $p_{ij}$  representa la probabilidad de que estado  $s_{t-1} = i$  venga seguido del estado  $s_t = j$ . Es decir,  $P(s_t = j | s_{t-1} = i) = p_{ij}$ . Este es el caso básico de las cadenas de Markov para dos estados, donde la ocurrencia de un estado elimina al otro o, dicho de otro modo, cada elemento de una columna es complementario del otro. Podemos entonces escribir una expresión equivalente a la anterior del siguiente modo:

$$\begin{bmatrix} p_{11} & p_{21} \\ 1 - p_{11} & 1 - p_{21} \end{bmatrix} \quad \text{Ec. 109}$$

Formalmente, se define una cadena de Markov de  $N$  estados como aquel proceso en el que la probabilidad de que un determinado estado  $s_t$  sea igual a un valor " $j$ ", comprendido entre 1 y  $N$ , sólo depende del valor que dicho estado haya tomado en el momento anterior. Dicho de otra forma:

$$P(s_t = j | s_{t-1} = i, s_{t-2} = k, \dots, s_{t-n} = z) = P(s_t = j | s_{t-1} = i) = \{p_{ij}\}_{i,j=1 \dots N} \quad \text{Ec. 110}$$

A esta expresión se la conoce como la ley básica de las cadenas de Markov. A la matriz que contiene todas las probabilidades para cada uno de los  $N$  estados posibles, se le conoce con el nombre de "matriz de transición"  $P$ :

$$\begin{bmatrix} p_{11} & p_{21} & \cdots & p_{N1} \\ p_{12} & p_{22} & \cdots & p_{N2} \\ \cdots & & \cdots & \cdots \\ p_{1N} & p_{2N} & \cdots & p_{NN} \end{bmatrix} \quad \text{Ec. 111}$$

(N x N)

Donde, por ejemplo, la fila 3 columna 4 muestra la probabilidad de que el estado 3 sea seguido del estado 4 (de un total de N estados diferentes).

### Relación de vectores autorregresivos en las cadenas de Markov

Si generamos una variable aleatoria, vector columna de orden (N x 1), que toma valor uno cuando  $s_t = j$  y cero en el resto de los casos, podríamos escribir, para todos los estados posibles, la matriz identidad:

$$\varepsilon_t = \begin{cases} (1,0,0,\dots,0)' & \text{si } s_t = 1 \\ (0,1,0,\dots,0)' & \text{si } s_t = 2 \\ \dots & \dots \\ (0,0,0,\dots,1)' & \text{si } s_t = N \end{cases} \quad \text{Ec. 112}$$

Cuando  $s_t = i$ , el vector  $\varepsilon_{t+1}$  tendrá probabilidad  $p_{ij}$  de tener un uno en su elemento "j" y cero en el resto, con lo cual, la esperanza condicional de este vector se puede escribir como:

$$E(\varepsilon_{t+1} | s_t = i) = \begin{bmatrix} p_{i1} \\ p_{i2} \\ \dots \\ p_{iN} \end{bmatrix} \quad \text{Ec. 113}$$

Para determinar ahora la esperanza condicional de  $\varepsilon_{t+1}$  ante un valor de  $\varepsilon_t$  cualquiera, podemos escribir:

$$E(\varepsilon_{t+1} | \varepsilon_t) = P \varepsilon_t \quad \text{Ec. 114}$$

que, debido a la ley básica enunciada en la expresión Ec. 110, es válida para toda la cadena, luego:

$$E(\varepsilon_{t+1} | \varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots) = P\varepsilon_t \quad \text{Ec. 115}$$

Si denominamos  $v_{t+1}$  a la diferencia entre el valor esperado y el valor real de  $\varepsilon_{t+1}$ , se puede escribir la típica forma de un proceso autorregresivo de primer orden para el caso que nos ocupa:

$$\varepsilon_{t+1} = P\varepsilon_t + v_{t+1} \quad \text{Ec. 116}$$

De cara a la obtención de valores de predicción en una Cadena de Markov así expresada, es fácil (a partir de la realización de sucesivas sustituciones del término  $\varepsilon_t$ ), llegar a la siguiente expresión:

$$\varepsilon_{t+m} = P^m \varepsilon_t + P^{m-1} v_{t+1} + P^{m-2} v_{t+2} + P^2 v_{t+m-2} + P v_{t+m-1} + v_{t+m} \quad \text{Ec. 117}$$

Dado que la esperanza de  $v_t$  es igual cero, la esperanza de la expresión anterior se puede reducir a:

$$E(\varepsilon_{t+m}) = P^m \varepsilon_t \quad \text{Ec. 118}$$

El vector  $\varepsilon_{t+m}$  tendrá valor uno cuando  $s_{t+m}=j$  y cero en el resto de los casos. En la expresión anterior, el elemento "j" del vector (N x 1) resultante del cálculo de la esperanza, representa la probabilidad de que  $s_{t+m}$  sea igual a "j" condicionada por el estado habido en "t". Si denominamos  $e_i$  al vector columna correspondiente al estado "i" (fila "i" de la matriz identidad), podemos escribir:

$$\begin{bmatrix} P(s_{t+m} = 1 | s_t = i) \\ P(s_{t+m} = 2 | s_t = i) \\ \dots \\ P(s_{t+m} = N | s_t = i) \end{bmatrix} = P^m e_i \quad \text{Ec. 119}$$



Con lo que se sigue que el valor de predicción para "m" periodos de una cadena de Markov se corresponde con la potencia "m" de la matriz de transición; es decir,  $P^m$ . Con ello, estamos diciendo que la probabilidad de que un estado "i" sea seguido de un estado "j" "m" periodos posteriores es igual a la fila "j", columna "i" de la matriz  $P^m$ .

### Ergodicidad de las cadenas de Markov

Dada la forma en la que se ha construido la matriz de transición P, en la que la suma por columnas es igual a uno, es seguro que este valor será precisamente uno de sus autovalores<sup>10</sup> (ya que una matriz y su traspuesta comparten los mismos autovalores).  $P' i = i$ , donde  $i$  es un vector de (N x 1) unos.

Se define un proceso como ergódico para una cadena de Markov de N estados como aquella que tiene todos sus autovalores dentro del círculo unitario, menos el que es uno. El vector de probabilidades ergódicas se puede escribir como:

$$P\pi = \pi \quad \text{Ec. 120}$$

Como el autovector estará normalizado (todos sus valores suman la unidad), se puede demostrar que:

$$\lim_{m \rightarrow \infty} P^m = \pi i' \quad \text{Ec. 121}$$

Con lo que queda asegurada la convergencia futura del proceso y, por tanto, su ergodicidad. De esta expresión, se deduce el valor esperado de  $\varepsilon_{t+m}$  converge en probabilidad a  $\pi i$ :

$$E(\varepsilon_{t+m} | \varepsilon_t, \varepsilon_{t-1}, \dots) = P^m \varepsilon_t \xrightarrow{p} \pi i' \varepsilon_t = \pi \quad \text{Ec. 122}$$

<sup>10</sup> Claramente, si tenemos una matriz A con un autovalor  $\lambda$  y un autovector asociado a ésta x, por definición:

ya que  $i' \varepsilon_t = 1$ . Con lo cual, el valor de predicción a largo plazo de  $\varepsilon_{t+m}$  viene dado por el vector de probabilidades ergódicas  $\pi$ , sin tener en cuenta el valor de este en el período  $t$ .

Además, el vector  $\pi$  representa las probabilidades marginales de cada uno de los  $N$  estados posibles<sup>11</sup>:  $E(\varepsilon_t) = \pi = P(s_t = j)$ .

En el caso en que los autovectores calculables de la matriz  $P$  sean distintos, Cox y Miller (1965) proponen una descomposición tipo Jordan para llegar al valor de  $P^m$ , obteniéndose el siguiente resultado:

$$P^m = T A^m T^{-1} \quad \text{Ec. 123}$$

donde  $T$  es una matriz que contiene cada uno de los autovectores de  $P$  y  $A$  es una matriz diagonal cuyos elementos se corresponden con los autovalores (como está elevada a " $m$ ", aparecerán cada uno de los autovalores elevados a esta cifra).

Para calcular las probabilidades ergódicas de una cadena de Markov con  $N$  estados se procede del siguiente modo:

Para un proceso de Markov con  $N$  estados, se busca un vector  $\pi$  tal que  $i' \pi = 1$  y que satisfaga la siguiente ecuación:

$$A \pi = e_{N+1}$$

donde  $A$  es una matriz de  $((N+1) \times N)$  de la forma:

---

$Ax = \lambda x$ . Dado que planteamos el caso en el que  $P'i = i$ , es obvio que el valor de  $\lambda$  igual a uno es una solución al sistema anterior.

<sup>11</sup> Si recurrimos a la fórmula vista en la Ec. 118, el valor de predicción para un período será  $E(\varepsilon_{t+1}) = PE(\varepsilon_t)$ . Planteado un proceso estacionario, el valor de la esperanza ha de ser constante. En esta expresión  $P$  es una matriz conocida (la matriz de transición) y, para que se cumpla la igualdad, es necesario entonces que podamos contar con un autovector de  $P$ , tal que  $P\pi = \pi$ , luego ha de ser que la  $E(\varepsilon_t) = \pi = P(s_t = j)$ . En definitiva,  $\pi$  es necesariamente el vector de probabilidades marginales del proceso  $E(\varepsilon_t)$  cuando este es estacionario.

$$A = \begin{bmatrix} I_N - P \\ i' \end{bmatrix}$$

con lo cual, el vector  $\pi$  buscado es la columna  $(N+1)$  de la matriz  $[A' A]^{-1} A'$

### Modelización de procesos con regímenes cambiantes

Partiendo de un proceso AR(1) del tipo:

$$y_t = c_{s_t} + \phi_{s_t} y_{t-1} + \varepsilon_t \quad \text{Ec. 124}$$

donde la perturbación aleatoria se distribuye como una  $N(0, \sigma^2)$  y es independiente de  $s_t$ , se pretende modelizar el estado  $s_t$  inobservable en "t" a partir de una cadena de Markov. Con este modelo se recogen posibles cambios de régimen permanentes o coyunturales y, además, admite la inclusión de otras variables explicativas para provocar estos cambios de régimen.

Agrupadas en un conjunto todas las posibles variables del modelo  $\mathcal{G}_t = (y_t', y_{t-1}', y_{t-m}', x_t', x_{t-1}', x_{t-m}')$ , donde  $y_t$  es un vector fila de  $(N \times 1)$  y  $x_t$  contiene todas las explicativas; la función de densidad condicional de  $y_t$ , condicionada por el estado  $s_t$  viene representada por:

$$f(y_t | s_t = j; x_t, \mathcal{G}_t; \alpha) \quad \text{Ec. 125}$$

donde  $\alpha$  contiene todos los parámetros desconocidos.

En el caso del modelo AR(1) representado en la Ec. 124,  $y_t$  es un vector de  $(1 \times 1)$ ,  $x_t$  es la variable explicativa constante y  $\alpha$ , contiene todos los parámetros posibles para los  $N$  estados teóricos:  $\alpha' = (c_1; c_2; \dots; c_N; \phi_1; \phi_2; \dots; \phi_N)$ . Para el caso en el que sólo hubiera dos estados posibles, la función de densidad de los estados 1 y 2 ( $\eta_t$ ) vendría definida por:

$$\eta_t = \begin{bmatrix} f(y_t | s_t = 1, y_{t-1}; \alpha) \\ f(y_t | s_t = 2, y_{t-1}; \alpha) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ \frac{-(y_t - c_1 - \phi_1 y_{t-1})^2}{2\sigma^2} \right\} \\ \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ \frac{-(y_t - c_2 - \phi_2 y_{t-1})^2}{2\sigma^2} \right\} \end{bmatrix} \quad \text{Ec. 126}$$

En esta situación, sólo se admite dependencia entre  $y_t$  y el estado contemporáneo a esta observación de la endógena ( $s_t$ ). Para dotar de más realismo a la especificación, se podría incorporar la importancia de  $s_{t-1}$  a partir de una Cadena de Markov, ahora de cuatro estados, como la siguiente:

$$s_t = 1 \quad \text{si} \quad s_t^* = 1 \text{ y } s_{t-1}^* = 1$$

$$s_t = 2 \quad \text{si} \quad s_t^* = 2 \text{ y } s_{t-1}^* = 1$$

$$s_t = 3 \quad \text{si} \quad s_t^* = 1 \text{ y } s_{t-1}^* = 2$$

$$s_t = 4 \quad \text{si} \quad s_t^* = 2 \text{ y } s_{t-1}^* = 2$$

La matriz de transición para esta cadena de Markov, denotando como  $p_{ij} = P(s_t = j | s_{t-1} = i)$ , sería la siguiente:

$$P = \begin{bmatrix} p_{11}^* & 0 & p_{11}^* & 0 \\ p_{12}^* & 0 & p_{12}^* & 0 \\ 0 & p_{21}^* & 0 & p_{21}^* \\ 0 & p_{22}^* & 0 & p_{22}^* \end{bmatrix} \quad \text{Ec. 127}$$

y la función de densidad correspondiente a los distintos estados sería:

$$\begin{aligned}
 f(y_t | y_{t-1}, s_t = 1, y_{t-1}; \alpha) &= \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{[(y_t - \mu_1 - \phi_1(y_{t-1} - \mu_1))]^2}{2\sigma^2} \right\} \\
 f(y_t | y_{t-1}, s_t = 2, y_{t-1}; \alpha) &= \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{[(y_t - \mu_2 - \phi_1(y_{t-1} - \mu_1))]^2}{2\sigma^2} \right\} \\
 f(y_t | y_{t-1}, s_t = 3, y_{t-1}; \alpha) &= \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{[(y_t - \mu_1 - \phi_1(y_{t-1} - \mu_2))]^2}{2\sigma^2} \right\} \\
 f(y_t | y_{t-1}, s_t = 4, y_{t-1}; \alpha) &= \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{[(y_t - \mu_2 - \phi_1(y_{t-1} - \mu_2))]^2}{2\sigma^2} \right\}
 \end{aligned}$$

Ec. 128

### Modelo ARCH de Régimen cambiante (*Switching ARCH*)

En diversos estudios en los que se han aplicado modelos tipo ARCH ya comentados anteriormente, se ha puesto de manifiesto la persistencia de la volatilidad en las series temporales. Algunos autores como Lamoureux y Lastrapes (1990) sugieren la posibilidad de que dicha persistencia sea debida a la existencia de cambios estructurales en los modelos utilizados. Redundando en este hecho, Perron (1989) argumenta que la persistencia en volatilidad, a veces entendida como la existencia de una raíz unitaria en el proceso, está ocultando el verdadero problema: parámetros no fijos en el modelo.

Hamilton y Susmel (1989) proponen un modelo de los conocidos como EGARCH o GARCH exponencial de Nelson (1990) con una peculiaridad añadida: se introduce en la especificación del proceso la dependencia de un determinado estado  $s_t$  inobservable, en el sentido de las Cadenas de Markov, estado que será el causante de que la media del proceso sea diferente en distintas submuestras.

Así, se define la probabilidad de que se dé un determinado estado  $s_t = j$ , condicionado por el resultado obtenido en el período anterior  $s_{t-1} = i$  del modo que se hizo en la Ec. 110, es decir:

$$P(s_t = j | s_{t-1} = i, s_{t-2} = k, \dots, s_{t-n} = z) = P(s_t = j | s_{t-1} = i) = \{p_{ij}\}_{i,j=1 \dots N}$$

Los diferentes estados posibles determinarán los parámetros de la función de distribución de un proceso estocástico  $\varepsilon_t$ , que se pretende modelizar con un ARCH, de modo que tendremos:

$$f(\varepsilon_t | s_t, s_{t-1}, \dots, s_{t-q}; \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}) \quad \text{Ec. 129}$$

En un modelo tipo ARCH, se podrían caracterizar distintos regímenes en la evolución de la varianza con una doble especificación sobre el comportamiento del proceso  $\varepsilon_t$  de la siguiente forma:

$$\varepsilon_t = \sqrt{g_{s_t}} \tilde{\varepsilon}_t \quad \text{Ec. 130}$$

$$\tilde{\varepsilon}_t = h_t^{1/2} v_t \quad \text{Ec. 131}$$

$$h_t = \alpha_0 + \alpha_1 \tilde{\varepsilon}_{t-1}^2 + \alpha_2 \tilde{\varepsilon}_{t-2}^2 + \dots + \alpha_q \tilde{\varepsilon}_{t-q}^2 + \theta d_{t-1} \tilde{\varepsilon}_{t-1}^2 \quad \text{Ec. 132}$$

donde,  $v_t \Rightarrow N(0,1)$  y  $d_{t-1}$  toma el valor 1 cuando  $\tilde{\varepsilon}_{t-1}^2 \leq 1$  y cero el resto de las veces. Con ello, la variable modelizada con un ARCH(q), tomará distintos valores en función de las diferentes realizaciones de los estados  $s_t$ .

De cara a representar los distintos multiplicadores  $g(\cdot)$  como cambios en la escala del proceso, se normaliza el proceso para órdenes de  $g_j \geq 2, 3, \dots, N$ . Es decir, despejando en la Ec. 130 tenemos:

$$\frac{\varepsilon_t^2}{g_{s_t}} = \tilde{\varepsilon}_t^2 \quad \text{Ec. 133}$$

lo que llevado a las Ec. 131 y a la Ec. 132, nos permite calcular la varianza del proceso  $\varepsilon_t$ , condicionada por los diferentes estados, como:

$$E(\varepsilon_t^2 | s_t, s_{t-1}, \dots, s_{t-q}, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}) = g_{s_t} \left[ (\alpha_0 + \alpha_1 (\varepsilon_{t-1}^2 / g_{s_{t-1}}) + \alpha_2 (\varepsilon_{t-2}^2 / g_{s_{t-2}}) + \dots + \alpha_q (\varepsilon_{t-q}^2 / g_{s_{t-q}}) + \vartheta d_{t-1} (\varepsilon_{t-1}^2 / g_{s_{t-1}})) \right] = \sigma_t^2(s_t, s_{t-1}, \dots, s_{t-q}) \quad \text{Ec. 134}$$

Si el valor estimado del término de apalancamiento fuera nulo, ( $\vartheta = 0$ ) el proceso existente se conoce con el nombre de SWARCH (N,q) o proceso ARCH(q) con Cadenas de Markov de N estados.

### Estimación y contraste

El logaritmo de la función de verosimilitud se maximizará de acuerdo a las siguientes restricciones:

- $g_1 = 1$
- $\sum_{j=1}^N p_{ij} = 1$  para  $i = 1, 2, \dots, N$  y  $0 \leq p_{ij} \leq 1$

El algoritmo de resolución se desarrolla del siguiente modo:

En el momento "t" del tiempo, los valores con los que se cuenta son<sup>12</sup>:

$$p(s_t, s_{t-1}, \dots, s_{t-q} | y_t, y_{t-1}, \dots, y_{-3}) \quad \text{Ec. 135}$$

donde cada uno de los N+1 números representados están multiplicados por  $p_{s_t, s_{t-1}}$  y por la función de densidad condicional  $f(y_{t+1} | s_{t+1}, s_t, s_{t-1}, \dots, s_{t-q}, y_{t+1} | y_t, y_{t-1}, \dots, y_{-3})$ . Estos sirven para obtener los N+2 valores diferentes de:

$$p(s_{t+1}, s_t, s_{t-1}, \dots, s_{t-q}, y_{t+1} | y_t, y_{t-1}, \dots, y_{-3}) \quad \text{Ec. 136}$$

\*

Si se entiende que el proceso es gaussiano, el cálculo anterior se realiza a partir de la función de máxima verosimilitud de la normal:

$$f(y_{t+1} | s_{t+1}, s_t, s_{t-1}, \dots, s_{t-q+1} | y_t, y_{t-1}, \dots, y_{t-q+1}) = \frac{1}{\sqrt{2\pi}\sigma_{t+1}(s_{t+1}, s_t, s_{t-1}, \dots, s_{t-q+1})} \exp\left(\frac{-(y_{t+1} - \alpha - \phi y_t)^2}{2\sigma_{t+1}^2(s_{t+1}, s_t, s_{t-1}, \dots, s_{t-q+1})}\right) \quad \text{Ec. 137}$$

donde la varianza  $\sigma_t^2(s_t, s_{t-1}, \dots, s_{t-q})$  viene dada tal y como se expresó en la Ec. 134, es decir:

$$E(\varepsilon_t^2 | s_t, s_{t-1}, \dots, s_{t-q}, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}) = g_{s_t}[(\alpha_0 + \alpha_1(\varepsilon_{t-1}^2 / g_{s_{t-1}}) + \alpha_2(\varepsilon_{t-2}^2 / g_{s_{t-2}}) + \dots + \alpha_q(\varepsilon_{t-q}^2 / g_{s_{t-q}}) + d_{t-1}(\varepsilon_{t-1}^2 / g_{s_{t-1}}))] = \sigma_t^2(s_t, s_{t-1}, \dots, s_{t-q}) \quad \text{Ec. 138}$$

donde  $d_{t-1}$  toma valor uno si la innovación en  $t+1$  es negativa y cero en el resto de los casos.

La función de densidad condicional de  $y_{t+1}$  es el resultado de la suma de los valores obtenidos en la Ec. 136, de tal forma que:

$$f(y_{t+1} | y_t, y_{t-1}, y_{t-2}) = \sum_{s_{t+1}=1}^N \sum_{s_t=1}^N \dots \sum_{s_{t-q}=1}^N p(s_{t+1}, s_t, s_{t-1}, \dots, s_{t-q}, y_{t+1} | y_t, y_{t-1}, \dots, y_{t-3}) \quad \text{Ec. 139}$$

de forma que se puede maximizar el logaritmo de la función máximo-verosímil, para las  $T$  observaciones existentes, escrito como:

<sup>12</sup> Los valores de partida para comenzar a iterar en la función  $p(s_0, s_{-1}, \dots, s_{-q} | y_0, y_{-1}, \dots, y_{-3})$  son iguales a las probabilidades ergódicas implícitas en la Cadena de Markov cuya forma de cálculo ya se expresó en un apartado anterior.



$$L = \sum_{t=1}^T \ln(f(y_t | y_{t-1}, y_{t-2}, \dots, y_{-3})) \quad \text{Ec. 140}$$

El algoritmo iterativo de resolución propuesto por Hamilton y Susmel es el de Kim's (1994).

De cara al contraste de la especificación propuesta y bajo el supuesto de que la especificación planteada se puede estimar, se propone la diferencia al cuadrado entre las varianzas condicional y constante del modelo propuesto:

$$MSE = T^{-1} \sum_{t=1}^T (\hat{\varepsilon}_t^2 - h_t)^2 \quad \text{Ec. 141}$$

Con ello se pretende juzgar la ganancia que se produce al estimar la varianza mediante este proceso condicional frente a la consideración de ésta como una constante. Si el valor de estas expresión es superior al de considerar que el modelo tiene varianza constante, estaríamos estimando, con los procesos GARCH, estructuras más ineficientes.

Suponiendo que el proceso  $\varepsilon_t$  fuera la perturbación aleatoria de, por ejemplo, un modelo AR(1) del siguiente tipo:  $z_t = \alpha_0 + \alpha_1 z_{t-1} + \varepsilon_t$ , para realizar dicha comparación, se calcularía el estadístico adicional:

$$MSE^{var\ cte} = T^{-1} \sum_{t=1}^T ((y_t - \bar{y})^2 - s^2)^2$$

$$s^2 = T^{-1} \sum_{t=1}^T (y_t - \bar{y})^2$$

Dado que en el proceso GARCH están interviniendo los momentos de cuarto orden, Hamilton y Susmel recalcan la pérdida de bondad de este contraste y proponen, en su lugar, el empleo de aquellos que señalaran West, Edison y Cho (1993):

- Diferencia en valor absoluto entre las varianzas condicional y constante del modelo

propuesto:  $MAE = T^{-1} \sum_{t=1}^T |\hat{\varepsilon}_t^2 - h_t| \quad \text{Ec. 142}$

- Diferencia al cuadrado entre los logaritmos de las varianzas condicional y constante

$$\text{del modelo propuesto: } [LE]^2 = T^{-1} \sum_{t=1}^T (\ln(\hat{\varepsilon}_t^2) - \ln(h_t))^2 \quad \text{Ec. 143}$$

- Diferencia al cuadrado entre los logaritmos de las varianzas condicional y constante

$$\text{del modelo: } LE = T^{-1} \sum_{t=1}^T |\ln(\hat{\varepsilon}_t^2) - \ln(h_t)| \quad \text{Ec. 144}$$

Para determinar el grado de "persistencia" de la volatilidad implícito en este modelo, se procede al siguiente desarrollo:

Partiendo del modelo propuesto en la Ec. 132:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 + \theta d_{t-1} \varepsilon_{t-1}^2 \quad \text{Ec. 145}$$

y escribiendo la expresión del valor condicional esperado para el momento  $t+m$  (empleando ya la notación GARCH), obtenemos:

$$E(h_{t+m}) \equiv E(\varepsilon_{t+m}^2 | \varepsilon_{t+m-1}, \varepsilon_{t+m-2}) = \alpha_0 + \alpha_1 \varepsilon_{t+m-1}^2 + \theta d_{t+m-1} \varepsilon_{t+m-1}^2 + \beta_1 h_{t+m-1} \quad \text{Ec. 146}$$

Dado que el último período histórico conocido se refiere al momento  $t$ , para obtener estos valores de predicción la condicionalidad impuesta se refiere a los valores de  $\varepsilon_t$  y anteriores. El valor de la esperanza condicional será entonces<sup>13</sup>:

$$E(\varepsilon_{t+m}^2 | \varepsilon_t, \varepsilon_{t-1}, \dots) = \alpha_0 + \alpha_1 E(\varepsilon_{t+m-1}^2 | \varepsilon_t, \varepsilon_{t-1}, \dots) + \theta E(d_{t+m-1} | \varepsilon_t, \varepsilon_{t-1}, \dots) E(\varepsilon_{t+m-1}^2 | \varepsilon_t, \varepsilon_{t-1}, \dots) + \beta_1 E(h_{t+m-1} | \varepsilon_t, \varepsilon_{t-1}, \dots)$$

La expresión anterior se puede resumir en:

<sup>13</sup> Obviamente, para la variable de apalancamiento  $d_{t+m-1}$  la construcción del modelo con un  $v_t$  simétrico exige que sea independiente de  $\varepsilon_t^2$ . Por otro lado, su valor condicionado será igual a ella misma.

$$\hat{h}_{t+m} \equiv E(\varepsilon_{t+m}^2) = \alpha_0 + (\alpha_1 + \beta_1 + \vartheta/2)\hat{h}_{t+m-1} = \alpha_0 + \lambda\hat{h}_{t+m-1} \quad \text{Ec. 147}$$

donde  $\lambda$  viene a representar la "persistencia" de los choques (*shocks*) de volatilidad que lleva implícita el modelo.

### Predicción

Para obtener los valores de predicción de un modelo SWARCH(N,q) se puede proceder del siguiente modo.

En primer lugar, el valor esperado en el período  $t+m$  para el proceso descrito según las Ec. 132, Ec. 131 y Ec. 130, sería el siguiente<sup>14</sup>:

$$\begin{aligned} E(\varepsilon_{t+m}^2 | s_t, s_{t-1}, \dots, s_{t-q+1}, \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}) = \\ E(g_{s_t} \tilde{\varepsilon}_{t+m}^2 | s_t, s_{t-1}, \dots, s_{t-q+1}, \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}) = \\ E(g_{s_t} | s_t, s_{t-1}, \dots, s_{t-q+1}) E(\tilde{\varepsilon}_{t+m}^2 | \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}) \end{aligned} \quad \text{Ec. 148}$$

Como el primer término del producto final de la anterior expresión es una cadena de Markov, se puede escribir la siguiente igualdad:

$$E(g_{s_t} | s_t, s_{t-1}, \dots, s_{t-q+1}) = \sum g_j \text{Prob}(s_{t+m} = j | s_t) \quad \text{Ec. 149}$$

El valor de predicción en  $t+m$  de la matriz de transición es igual a "m" veces el producto de la matriz P, tal y como se detallaba en la expresión de la Ec. 119:

<sup>14</sup> Teniendo en cuenta que  $s_t$  es independiente de  $v_t$  y de  $\tilde{\varepsilon}_t$  en cualquier momento del tiempo.

$$\begin{bmatrix} P(s_{t+m} = 1 | s_t = i) \\ P(s_{t+m} = 2 | s_t = i) \\ \dots \\ P(s_{t+m} = N | s_t = i) \end{bmatrix} = P^m e_i \quad \text{Ec. 150}$$

donde  $e_i$  denotaba la columna  $i$ -ésima de la matriz identidad de orden  $(N \times N)$ .

Si se recogen en un vector columna de  $(N \times 1)$  los diferentes valores de la función  $g$ , se puede escribir:

$$\begin{aligned} g' &= (g_1, g_2, \dots, g_N) \\ E(g_{s_{t+m}} | s_t = i) &= g' P^m e_i \end{aligned} \quad \text{Ec. 151}$$

El segundo término de la ecuación Ec. 148 es un modelo ARCH en logaritmos de orden "q", que se puede escribir como:

$$\begin{aligned} E(\tilde{\varepsilon}_{t+m}^2 | \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}) &= \\ &= \alpha_0 + \alpha_1 \tilde{\varepsilon}_t^2 + \alpha_1 \tilde{\varepsilon}_{t-1}^2 + \alpha_q \tilde{\varepsilon}_{t-q+1}^2 + \theta \tilde{\varepsilon}_t^2 \quad \text{para } m = 1 \\ &= \alpha_0 + (\alpha_1 + \theta/2) \tilde{h}_{t+m-1|t} + \alpha_2 \tilde{h}_{t+m-2|t} + \dots + \alpha_q \tilde{h}_{t+m-q|t} \quad \text{para } m = 2, 3, \dots \end{aligned} \quad \text{Ec. 152}$$

donde,

$$\begin{aligned} h_{\tau|t} &= \tilde{\varepsilon}_t^2 \quad \text{para } \tau \leq t \\ &= E(\tilde{\varepsilon}_\tau^2 | \tilde{\varepsilon}_t^2, \tilde{\varepsilon}_{t-1}^2, \dots) \quad \text{para } \tau > t \end{aligned}$$

Los valores de  $h_{\tau|t}$  cuando  $\tau = t+2, t+3, \dots$  se calculan iterando en la expresión Ec. 152.

Ya que en el modelo propuesto  $\tilde{\varepsilon}_t = \varepsilon_t / \sqrt{g_{s_t}}$ , la predicción de la Ec. 152 está en función

de  $\varepsilon_t, \varepsilon_{t-1}, \dots$ , con lo que la Ec. 148

se pueden escribir ahora como:

$$\begin{aligned} E(\varepsilon_{t+m}^2 | s_t, s_{t-1}, \dots, s_{t-q+1}, \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}) &= \\ = E(g_{s_t} | s_t, s_{t-1}, \dots, s_{t-q+1}) E(\tilde{\varepsilon}_{t+m}^2 | \tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-1}, \dots, \tilde{\varepsilon}_{t-q+1}) &= \text{Ec. 153} \\ = g' P^m e_{s_t} \tilde{h}_{t+m|t}(s_t, s_{t-1}, \dots, s_{t-q+1}, \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) \end{aligned}$$

En la realidad, los diferentes estados o valores de  $s_t, s_{t-1}$ , etc. no son conocidos. Aún así, y gracias a la ley de las expectativas iteradas, se puede escribir:

$$\begin{aligned} h_{t+m|t} &= E(\varepsilon_{t+m}^2 | \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) = \\ &= \sum_{s_t=1}^N \sum_{s_{t-1}=1}^N \dots \sum_{s_{t-q+1}=1}^N \left[ (N(s_t, s_{t-1}, s_{t-q+1}, \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1})) (p(s_t, s_{t-1}, s_{t-q+1} | z_t, z_{t-1}, \dots)) \right] \text{Ec. 154} \end{aligned}$$

El resultado obtenido es una suma ponderada de las predicciones condicionales hasta el período "m".

Con el modelo propuesto se asume que si el fenómeno ha estado en el régimen 1 en los último "q" períodos y la  $p_{11}$  es cercana a la unidad (la probabilidad de que el estado 1 venga precedido de otro de igual valor)

- si la innovación el último período "t" ha sido pequeña, la predicción sería, básicamente, igual:

$$E(\varepsilon_{t+1}^2 | \varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-q+1}) = \alpha_0 + \alpha_1 (\varepsilon_t^2 / g_1) + \alpha_2 (\varepsilon_{t-1}^2 / g_1) + \alpha_q (\varepsilon_{t-q+1}^2 / g_1) + \partial d_t (\varepsilon_t^2 / g_1)$$

Con lo que el efecto marginal de  $\varepsilon_t^2$  vendría dado por  $(\alpha_1 + \partial d_t) / g_1$

- si la innovación del último período fuera grande, el efecto marginal vendría dado por  $(\alpha_t + \beta d_t)/g_2$ .

En definitiva, la volatilidad admite diversas formas no lineales que se construyen a partir de la inferencia sobre el actual estado de la misma en el período "t".

Gourieroux y Montfort (1992) definen una variante de este tipo de modelo a la que dan el nombre de *Qualitative Threshold ARCH* (Modelos ARCH de umbral cualitativo). En su estudio, se determina con sencillez la forma en la que se construyen las funciones de autocorrelación total y parcial para identificar la situación propuesta por estos autores, generando un modelo apto para la introducción de más variables explicativas exógenas en la especificación de la varianza.

## 5.5. Modelos Asimétricos potenciales tipo ARCH (AP-ARCH)

### Introducción

En el año 1993, Zhanxin Ding, con la colaboración de Granger y Engle propone una nueva expresión generalizada en la familia ARCH cuyo principal objetivo es recoger dos circunstancias comunes en las series financieras:

- Además de existir una importante autocorrelación entre los valores de la perturbación aleatoria en términos absolutos, dicha perturbación (en valor absoluto) elevada a una potencia " $d$ " muestra aún mayores autocorrelaciones, incluso con un largo desfase temporal (la memoria temporal es muy amplia en estos procesos).
- No parece existir una razón clara para admitir a priori la relación lineal impuesta en todos los modelos tipo ARCH en la especificación de la varianza condicional en función de los residuos retardados al cuadrado o de la desviación típica retardada.
- Dado que la autocorrelación observable en las series financieras parece ser mayor en valores cercanos a uno (ligemente mayores que uno, generalmente) no parece idóneo emplear como variable a modelizar la varianza o la desviación típica en los modelos tipo ARCH, sino un cierto valor de la desviación típica elevada a una potencia " $\delta$ " por determinar (aquella con valores de autocorrelación más grandes en la variable aleatoria absoluta elevada a esa " $\delta$ ").

El modelo A-PARCH aquí propuesto acoge, en su especificación, como casos especiales los modelos tipo ARCH ya comentados previamente, aunque dota a la especificación de la capacidad de recoger otro tipo de relaciones no lineales entre la varianza condicional y sus factores explicativos.

### Especificación

El modelo A-PARCH definido por Ding (1993), usando la misma terminología empleada para el resto de los casos, se formularía del siguiente modo:

$$\begin{aligned}\varepsilon_t &= \eta_t h_t^{1/2} \\ \varepsilon_t | \psi_{t-1} &\rightarrow N(0, h_t) \\ h_t^\delta &= \alpha_0 + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^p \beta_j h_{t-j}^\delta \\ \alpha_0 &> 0; \delta \geq 0 \\ \alpha_i &\geq 0, i = 1, \dots, q \\ \beta_j &\geq 0, j = 1, \dots, p\end{aligned}$$

Obsérvese que la propuesta de Ding se centra en admitir una modelización de la desviación típica elevada no necesariamente a dos. Es decir, no se impone la modelización de la varianza como un proceso condicional autorregresivo, sino una cierta potencia de la desviación típica (aquella en la que la autocorrelación del proceso aleatorio en valor absoluto es máxima). En este modelo están implícitas dos transformaciones potenciales tipo Box-Cox:

- Sobre la desviación típica (en función del valor de  $\delta$ ); con la que se recoge el coeficiente de elasticidad de sustitución (CES) de un modelo tipo Arrow (1961).
- Sobre el valor absoluto asimétrico de los residuos; con la que se recoge el "efecto apalancamiento" (*leverage*) presente en los mercados financieros; es decir, la respuesta distinta ante "shocks" previos según el signo de estos.

Tal y como se ha comentado, esta nueva presentación de los modelos A-PARCH, acoge como casos especiales otros ya comunes<sup>15</sup>:

1. Modelo ARCH de Engle (1982): cuando  $\delta=2$ ,  $\gamma_i=0$  y  $\beta_j=0$
2. Modelo GARCH (p,q) de Bollerslev (1986): cuando  $\delta=2$ ,  $\gamma_i=0$  y  $\beta_j \neq 0$
3. Modelo GARCH de Taylor/Schwert (1986) en desviaciones a la media: cuando  $\delta = 1$ ,  $\gamma_i = 0$  y  $\beta_j \neq 0$ .
4. Modelo GJR (1989): cuando  $\delta=2$ ,  $0 < \gamma_i < 1$  y  $\beta_j \neq 0$



5. Modelo TARCH de Zakoian (1991): cuando  $\delta = 1$ ,  $\gamma_i = 0$  y  $\beta_j \neq 0$ .
6. Modelo NARCH de Higgins y Bera (1990): cuando  $\delta^* = \delta/2$ ,  $\gamma_i = 0$  y  $\beta_j \neq 0$ .
7. Modelo log-ARCH de Geweke y Pantula (1986): cuando  $\delta \rightarrow 0$ .

Como demuestra el autor en el apéndice B de su artículo de 1993, si se asume una distribución condicional normal del proceso aleatorio, para la existencia de los momentos de orden  $E(h_t^\delta)$  y  $E|\varepsilon_t|^\delta$  es necesario que:

$$\frac{1}{\sqrt{2\pi}} \sum_{i=1}^p \alpha_i [(1+\gamma_i)^\delta + (1-\gamma_i)^\delta] 2^{\frac{\delta-1}{2}} \Gamma \frac{\delta+1}{2} + \sum_{j=1}^q \beta_j < 1$$

Lo que se define como la condición de estacionariedad del proceso.

### Estimación y contraste

Ding advierte la conveniencia de aplicar una potencia  $\delta$  distinta de uno o de dos (como en los modelos en desviaciones típicas o en varianza) a partir de un estudio de Montecarlo realizado sobre un amplio conjunto de series financieras de diversos países. De su estudio, cabe concluir dos cosas:

- La autocorrelación en las potencias alrededor de uno del valor absoluto del proceso aleatorio mantienen mayor autocorrelación para cualquier retardo que las de cualquier otra potencia. En su caso, el valor de  $\delta$  óptimo parece ser el de 1,43.
- La observación anterior no varía aunque lo haga la frecuencia de las series utilizadas. Es decir; aunque las variables sean diarias, semanales, mensuales, etc. el valor de la potencia alrededor de uno es el que muestra mayor

---

<sup>15</sup> Para más detalles en la transformación del caso general a cada uno de los siete específicos comentados ver Ding (1993), Apéndice A.

autocorrelación muestral, siguiéndose una distribución monotónica decreciente a ambas colas desde este máximo absoluto.

Para la estimación del modelo, Ding propone emplear el algoritmo de resolución de Berndt-Hall-Hall-Hausman en la función normal de las características del proceso antes descrito.

Para determinar la utilización de un valor de la potencia distinto a, por ejemplo, uno, como en el caso de un modelo ARCH en media, o dos, como en un GARCH(p,q) del tipo definido por Bollerslev (1986), Ding propone realizar un contraste tipo ratio de verosimilitud o Breusch Pagan del tipo:

$$2(l_1 - l_0) \rightarrow \chi^2_2$$

$l_0$ : el logaritmo de máxima-verosimilitud bajo la hipótesis nula de que el modelo GARCH(p,q) es el correcto.

$l_1$ : el logaritmo de máxima-verosimilitud bajo la hipótesis alternativa de que el modelo A-PARCH es el correcto.

En el caso en que la diferencia propuesta se distribuya según el valor teórico (una chi-cuadrado con dos grados de libertad), se admitiría la hipótesis nula; es decir, se admitiría que el proceso de autocorrelación condicional de la varianza se rige según un GARCH(p,q). En caso contrario, se admitiría que el proceso es realmente un A-PARCH.

## **6. MODELOS DE FIJACIÓN DE PRECIOS DE ACTIVOS**

## 6.1 Modelos unifactoriales de fijación de precios de cartera. Modelos CAPM (Capital Asset Pricing Models)

### Introducción

Los modelos de *Asset Pricing* parten del presupuesto de que los agentes económicos maximizan su función de utilidad bajo la existencia de la incertidumbre sobre los sucesos futuros. Fijada una rentabilidad esperada de la inversión, los agentes intentarán determinar su posición en cada uno de los activos existentes en el mercado eligiendo, a partir de la información disponible, la combinación de éstos (cartera de valores o *portfolio*) con mínima volatilidad (*varianza condicional*) y máxima rentabilidad (*esperanza condicional*).

Markowitz (1959) define el problema de selección de la cartera de valores para el inversor en términos de rentabilidad esperada y volatilidad. Argumenta este autor (cita literal) que "los inversores optimizarán su utilidad manteniendo una cartera de activos con máxima rentabilidad esperada, dado un nivel mínimo de varianza". En definitiva, la CAPM se asienta en el supuesto de que la cartera de valores de mercado es una combinación de activos eficiente en media y varianza.

Sharpe (1964) y Lintner (1965) recogen las ideas de Markowitz definiendo un modelo como sigue:

$$E(R_i) = R_f + \beta_{im} (E(R_m) - R_f)$$
$$\beta_{im} = \frac{\text{cov}(R_i, R_m)}{\text{var}(R_m)}$$

Ec. 155

donde:

$R_i$ : rentabilidad del activo "i"

$R_m$ : rentabilidad media ponderada de los valores existentes en el mercado

$R_f$ : rentabilidad del activo sin riesgo

Esta teoría se describe con frecuencia en términos de "prima de rentabilidad sobre el activo sin riesgo", es decir, buscando la explicación del exceso de rentabilidad de un activo con riesgo sobre otro sin él:

$$\begin{aligned} E(Z_i) &= \beta_{im} (E(Z_m)) \\ \beta_{im} &= \frac{\text{cov}(Z_i, Z_m)}{\text{var}(Z_m)} \quad \text{Ec. 156} \\ Z_i &= R_i - R_f \\ Z_m &= R_m - R_f \end{aligned}$$

El significado financiero del parámetro  $\beta_{im}$  viene a representar, en esta expresión, la sensibilidad de la rentabilidad de la cartera de valores ante la volatilidad general del mercado. A esta expresión se la conoce con el nombre de "beta de inversión o beta financiera" y, en la medida en la que sea mayor que uno, hablaremos de inversiones relativamente arriesgadas, mientras que, en el caso contrario, hablaremos de una cartera menos sensible a los movimientos del mercado.

La aplicación de Sharpe-Lintner en esta segunda expresión tiene tres implicaciones:

- el valor de la constante es nulo,
- el parámetro  $\beta$  recoge completamente la variación de la rentabilidad esperada en el corte transversal,
- la prima de rentabilidad sobre el activo sin riesgo es positiva:  $Z_i = R_i - R_f > 0$ .

Black (1972) planteó una variante sobre este modelo CAPM general: el caso en el que no hubiera ningún activo sin riesgo. En su exposición, emplea los valores de rentabilidades reales, no nominales. En esta situación, el valor esperado de la rentabilidad del activo "i" quedaría explicado como sigue:

$$R_i = E(R_{0m}) + \beta_{im} (E(R_m) - E(R_{0m})) \quad \text{Ec. 157}$$

donde:

$R_{0m}$  es la rentabilidad de la cartera de valores "beta-cero" asociada al mercado "m". Esta composición de cartera se define como aquella que tiene la mínima varianza de entre todas las posibles carteras incorrelacionadas con "m", ya que el valor del parámetro era y es:

$$\beta_{im} = \frac{\text{cov}(R_i, R_m)}{\text{var}(R_m)} \quad \text{Ec. 158}$$

Siguiendo la interpretación común del parámetro beta ya comentado anteriormente, este activo teórico con rentabilidad  $R_{0m}$  sería el menos expuesto o sensible a variaciones en el resto del mercado.

La definición del modelo de Black se puede entender como el caso restringido del modelo más general del tipo  $E(R_i) = \alpha_{im} + \beta_{im} E(R_m)$ , donde lo que implicaría el modelo de Black sería que el término independiente fuera igual a:

$$\alpha_{im} = (1 - \beta_{im}) E(R_{0m}) \quad \text{Ec. 159}$$

Hasta aquí se han tratado de una serie de identidades matemáticas que no comportan más que un período de realización. De cara a la obtención de un modelo econométrico y la consiguiente inclusión del factor tiempo, se puede suponer que el comportamiento de las series temporales de las rentabilidades se distribuyen como una normal. Con ello, podemos estimar consistentemente el modelo:

$$Z_{it} = \alpha_{im} + \beta_{im} Z_{mt} + \varepsilon_{it} \quad \text{Ec. 160}$$

donde:

$Z_{it}$ : prima de riesgo del activo "i" (diferencia entre la rentabilidad del activo "i" y el activo sin riesgo" del mercado).

$Z_{mt}$ : prima de riesgo del total del mercado (por ejemplo, un índice general de la bolsa menos el valor de la rentabilidad de un activo público).

### Relaciones matemáticas e identidades implícitas del modelo para un solo período

De cara a cuantificar matemáticamente el equilibrio según la situación planteada en el modelo Sharpe-Lintner, se hacen las siguientes definiciones:

Sean:

$N$ : el número de activos con riesgo en el mercado cuyo vector de medias de rentabilidad es  $\mu$  y su covarianza  $\Omega$ .

$w_a$ : vector ( $N \times 1$ ) de pesos de cada activo en la cartera de valores "a" siendo la suma igual a uno  $\sum_{i=1}^N w_i = 1$

La rentabilidad media de la cartera de valores "a" será:  $\mu_a = w_a' \mu$  y su varianza  $\sigma_a^2 = w_a' \Omega w_a$ ,

La covarianza entre dos carteras de activos a y b será:  $w_a' \Omega w_b$ .

Según estas expresiones, la cartera de activos óptima será aquella que, dada una rentabilidad  $\mu_p$  minimice la expresión con restricciones:

$$\begin{aligned} \min_w \quad & w' \Omega w \\ & w' \mu = \mu_p \\ & w' i = 1 \end{aligned} \quad \text{Ec. 161}$$

que, escrita en forma de lagrangiano sería:

$$L = w' \Omega w - \delta_1 (\mu_p - w' \mu) - \delta_2 (1 - w' i) \quad \text{Ec. 162}$$

Calculando las derivadas parciales respecto a los parámetros presentados, obtenemos el siguiente sistema de ecuaciones con solución para cada valor dado de la rentabilidad esperada  $\mu_p$ :

$$\begin{aligned}\frac{\partial L}{\partial w} &= 2\Omega w - \delta_1 \mu - \delta_2 i = 0 \\ \frac{\partial L}{\partial \delta_1} &= \mu_p - w' \mu = 0 \\ \frac{\partial L}{\partial \delta_2} &= 1 - w' i = 0\end{aligned}\quad \text{Ec. 163}$$

Realizando una serie de sustituciones para lograr una solución inmediata al sistema, podemos escribir las tres condiciones anteriores en una sola del siguiente modo:

$$w_p = g + h\mu_p \quad \text{Ec. 164}$$

donde:

$$g = \frac{1}{D} [B(\Omega^{-1}i) - A(\Omega^{-1}\mu)] \quad \text{vector de (N x 1)}$$

$$h = \frac{1}{D} [C(\Omega^{-1}\mu) - A(\Omega^{-1}i)] \quad \text{vector de (N x 1)}$$

$$A = i' \Omega^{-1} \mu$$

$$B = \mu' \Omega^{-1} \mu$$

$$C = i' \Omega^{-1} i$$

$$D = BC - A^2$$

En definitiva, con la expresión de la Ec. 72 obtenemos un resultado matemático de asignación de proporción de cada activo en una cartera a la cual se le pide una rentabilidad prefijada  $\mu_p$ .

Obviamente, la solución puede no ser única (pueden coexistir varias carteras distintas con varianza mínima e igual rentabilidad). Para cada cartera de valores "p" con mínima



varianza, existe una cartera con covarianza cero con esta a la que se le llama habitualmente la cartera beta-cero.

A partir de las expresiones anteriores, se puede calcular la cartera de valores con mínima varianza posible, a la que llamaremos "g", y definir su rentabilidad media, su varianza y la matriz de pesos de cada activo del siguiente modo:

$$\begin{aligned}w_g &= \frac{1}{C} \Omega^{-1} i \\ \mu_g &= \frac{A}{C} \\ \sigma_g^2 &= \frac{1}{C}\end{aligned} \quad \text{Ec. 165}$$

### Estimación y contraste del modelo CAPM

Para la estimación del modelo CAPM se utiliza la expresión de la Ec. 160, según la cual definíamos el exceso de rentabilidad de un activo "i" respecto al activo sin riesgo como:

$$Z_i = \alpha + \beta Z_{m_i} + \varepsilon_i \quad \text{Ec. 166}$$

donde:

$$\begin{aligned}E(\varepsilon_i) &= 0 \\ E(\varepsilon_i, \varepsilon_i) &= \Sigma \\ E(Z_{m_i}) &= \mu_m \\ E(Z_{m_i} - \mu_m)^2 &= \sigma_m^2 \\ \text{Cov}(Z_{m_i}; \varepsilon_i) &= 0\end{aligned}$$

El supuesto intrínseco de la CAPM, tal y como ya se ha comentado y expresado en la Ec. 160, es que el vector de parámetros independientes tiene valor real igual a cero ( $\alpha=0$ ) y que la  $\beta$  es mayor que cero (prima de riesgo positiva).

Asumiendo normalidad en todas las variables aleatorias consideradas en la expresión anterior, se pueden estimar los parámetros de los  $N$  activos por máxima-verosimilitud, cuya expresión del logaritmo de la función de densidad conjunta será:

$$L = -\frac{NT}{2} \log(2\pi) - \frac{T}{2} \log |\Sigma| - \frac{1}{2} \sum_{i=1}^T (Z_i - \alpha - \beta Z_{mi})' \Sigma^{-1} (Z_i - \alpha - \beta Z_{mi}) \quad \text{Ec. 167}$$

La solución por derivadas parciales sería:

$$\begin{aligned} \frac{\partial L}{\partial \alpha} &= \Sigma^{-1} \left[ \sum_{i=1}^T (Z_i - \alpha - \beta Z_{mi}) \right] \\ \frac{\partial L}{\partial \beta} &= \Sigma^{-1} \left[ \sum_{i=1}^T (Z_i - \alpha - \beta Z_{mi}) Z_{mi} \right] \\ \frac{\partial L}{\partial \Sigma} &= -\frac{T}{2} \Sigma^{-1} + \frac{1}{2} \Sigma^{-1} \left[ \sum_{i=1}^T (Z_i - \alpha - \beta Z_{mi})(Z_i - \alpha - \beta Z_{mi})' \right] \Sigma^{-1} \end{aligned} \quad \text{Ec. 168}$$

Para la estimación de la matriz de varianzas covarianzas, es habitual el empleo de la matriz de información de Fisher; es decir, de la negativa de la esperanza de la derivada de segundo orden del logaritmo de máxima-verosimilitud respecto al vector de parámetros.

El resultado permite describir los dos primeros momentos y las funciones de distribución como<sup>16</sup>:

$$\begin{aligned} \hat{\alpha} &\Rightarrow N\left(\alpha; \frac{1}{T} \left[ 1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2} \right] \Sigma\right) \\ \hat{\beta} &\Rightarrow N\left(\beta; \frac{1}{T} \left[ 1 + \frac{1}{\hat{\sigma}_m^2} \right] \Sigma\right) \\ T\hat{\Sigma} &\Rightarrow W(T-2; \Sigma) \end{aligned} \quad \text{Ec. 169}$$

donde los estimadores consistentes de varianza y media de  $Z_{mi}$  son:

<sup>16</sup> La notación  $W$  se refiere a la distribución de Wishart con  $T-2$  grados de libertad y covarianza  $\Sigma$ , que reproduce una distribución chi-cuadrado para el caso multivariante.

$$\begin{aligned}\hat{\mu}_m &= \frac{1}{T} \sum_{i=1}^T Z_{mi} \\ \hat{\sigma}_m^2 &= \frac{1}{T} \sum_{i=1}^T (Z_{mi} - \hat{\mu}_m)^2\end{aligned}\quad \text{Ec. 170}$$

Para contrastar la hipótesis nula  $H_0(\alpha = 0)$  se construye un estadístico tipo Wald cuya expresión se demuestra que es la siguiente:

$$J_0 = \hat{\alpha}' (\text{var}(\hat{\alpha}))^{-1} \hat{\alpha} = T \left( 1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2} \right)^{-1} \hat{\alpha}' \Sigma^{-1} \hat{\alpha} \quad \text{Ec. 171}$$

Si se acepta la hipótesis nula,  $J_0$  se distribuirá como una  $\chi^2$  con N grados de libertad.

En esta expresión, el único valor desconocido es el de  $\Sigma$ , que puede sustituirse por el valor de su estimación por máxima verosimilitud que se exponía en la Ec. 168.

Dado que este test no puede ser aplicado en otro caso diferente al de muestras temporales muy largas, MacKinlay (1987) y Gibbons, Ross y Shanken (1989) proponen otro contraste ahora sí utilizable con muestras pequeñas. Su definición matemática es la siguiente:

$$J_1 = \frac{T-N-1}{N} \left( \left[ 1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2} \right]^{-1} \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} \right) \quad \text{Ec. 172}$$

con esta expresión se contrasta la hipótesis nula  $H_0(\alpha = 0)$  que, en caso de cumplirse, haría que el resultado de la Ec. 172 estuviera dentro de los márgenes de una F de Snedecor con N; T-N-1 grados de libertad.

### Aplicación al corte transversal

En muchas ocasiones la contrastación de esta teoría en términos de corte transversal resulta de especial interés para descubrir relaciones entre los distintos activos y carteras para el mismo momento del tiempo. Fama y Macbeth (1973) describieron un procedimiento para aplicar la CAPM a diversos activos contemplados estos en el mismo momento del tiempo. Los autores referidos proponen la siguiente regresión:

$$Z_t = \gamma_{0t}i + \gamma_{1t}\beta_m + \eta_t \text{ Ec. 173}$$

donde:

$Z_t$ : es un vector ( $N \times 1$ ) de  $N$  rentabilidades por encima de la libre de riesgo de activos diferentes evaluadas en el momento  $t$ .

$\gamma_{0t}$  y  $\gamma_{1t}$ : parámetros estimados en  $N$  regresiones diferentes para toda la extensión temporal de la que se dispone y usando MCO.

$\beta_m$ : vector de ( $N \times 1$ ) parámetros beta diferentes estimados por la CAPM comentada hasta el momento.

$\eta_t$ : perturbación aleatoria idénticamente distribuida como una normal y con las características habituales.

Fama y MacBeth proponen dos etapas en la estimación de este modelo:

1. Se estima la Ec. 173 de forma independiente para cada uno de los excesos de rentabilidad ( $N$  ecuaciones diferentes) y con una muestra de  $T$  observaciones temporales distintas.
2. Se analizan los valores estimados de los parámetros. Para ello, se define  $\gamma_0 = E(\gamma_{0t})$  y  $\gamma_1 = E(\gamma_{1t})$ . A continuación se contrastan las implicaciones de la CAPM; es decir, que el término constante es nulo y que el parámetro  $\gamma_1 > 0$ . Para contrastar este hecho, y partiendo de la distribución normal de la perturbación aleatoria y, por tanto, de los parámetros estimados por MCO, se emplea el clásico *test t-student* del siguiente modo:

$$t_{T-1}^{stad}(\hat{\gamma}_j) = \frac{\hat{\gamma}_j}{\hat{\sigma}_{\hat{\gamma}_j}} \Rightarrow t_{T-1} \text{ Ec. 174}$$

$H_0(\gamma_j=0)$

Estimados como:

$$\hat{\gamma}_j = \frac{1}{T} \sum_{t=1}^T \hat{\gamma}_{jt}$$

$$\sigma_{\hat{\gamma}_j}^2 = \frac{1}{T(T-1)} \sum (\hat{\gamma}_{jt} - \hat{\gamma}_j)^2$$

Ec. 175

Con frecuencia esta forma de contrastar la CAPM se emplea para incorporar a esta teoría variables explicativas adicionales a las comentadas hasta ahora. Por ejemplo, es sencillo seguir la misma dinámica aquí expuesta para contrastar si el tamaño de la cartera sirve para explicar la rentabilidad de esta, añadiendo a la regresión de la Ec. 173 una variable explicativa más ( $\zeta$ ) y contrastando, a partir de la t-student, si su parámetro a la hora de explicar  $Z_t$  es nulo o no.

A esta teoría se le hace una crítica fundamental. dado que el valor de las "betas" no es conocido, es necesario estimarlas para la aplicación del procedimiento descrito. Con ello se incluye una fuerte fuente de error en el modelo y un término aleatorio en los regresores, con todos los problemas que la economía tradicional describe en este caso en la estimación MCO.

## 6.2 Modelos multifactoriales de fijación de precios de cartera: Teoría del Arbitraje de precios (Arbitrage Pricing Theory - APT -)

### Introducción

La APT es introducida por Ross (1976) buscando una alternativa a la CAPM, dado los problemas de aplicación que esta segunda sufre por lo muy limitado de sus asunciones sobre las variables explicativas del modelo propuesto. En la APT se van a considerar múltiples causantes o factores de riesgo para determinar el precio de un activo. La principal aportación a su manejabilidad es que se va a establecer un sistema en el que no es necesario conocer de antemano la cartera de valores del total del mercado.

La asunción principal de esta teoría es que la rentabilidad de un activo "i" se puede explicar en función de un término constante y de una serie de factores, modificados por un parámetro de "sensibilidad" del activo frente a ellos y, finalmente, de un término aleatorio y desconocido. Matemáticamente su especificación sería la siguiente:

$$\begin{aligned} R_i &= a_i + b_i f + \varepsilon_i \\ E(\varepsilon_i | f) &= 0 \\ E(\varepsilon_i^2) &= \sigma_i^2 < \sigma^2 < \infty \end{aligned} \quad \text{Ec. 176}$$

- $R_i$ : rentabilidad del activo "i" (1 x 1)  
 $a_i$ : constante  
 $b_i$ : sensibilidad ante el factor (k x 1)  
 $f$ : vector de realizaciones comunes del factor (k x 1)

Para el caso en que se consideren N activos de forma simultánea, la especificación matricial del sistema propuesto sería la siguiente:

$$\begin{aligned} R &= a + Bf + \varepsilon \\ E(\varepsilon | f) &= 0 \\ E(\varepsilon \varepsilon') &= \Sigma \end{aligned} \quad \text{Ec. 177}$$

donde,

- $R$ : vector (N x 1) de las distintas rentabilidades de los activos  $R = [R_1, R_2, \dots, R_N]'$   
 $A$ : vector (N x 1) de términos constantes para cada activo  $a = [a_1, a_2, \dots, a_N]'$   
 $B$ : matriz (N x K) de sensibilidades del tipo  $B = [b_1, b_2, \dots, b_K]'$   
 $\varepsilon$ : vector (N x 1) de perturbaciones aleatorias  $\varepsilon = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N]'$

Dada esta situación, Ross afirma que la ausencia de arbitraje en economías suficientemente desarrolladas (mercados amplios y profundos) supone que la rentabilidad esperada de un activo sea igual a una constante (el parámetro del modelo beta-cero expresado en la sección anterior) más una cierta prima de riesgo. Matemáticamente, se podría escribir esta relación como:

$$\mu = i\lambda_0 + B\lambda_K \quad \text{Ec. 178}$$

donde,

$\mu$ : vector de (N x 1) rentabilidades esperadas de los activos

$\lambda_0$ : parámetro beta-cero

$B\lambda_K$ : lambda es un vector (K x 1) de prima de riesgo, modificada para cada activo "i" por un parámetro  $b_i$  de sensibilidad frente al factor.

Obviamente, la cuestión fundamental en los modelos APT es la determinación de los factores explicativos que dan lugar a esta identidad.

### Determinación de los factores

Desde el punto de vista teórico, la selección de los factores a incluir como explicativas de la rentabilidad de los activos se centra en dos grandes campos: el macroeconómico y el microeconómico.

Las principales variables empleadas como determinantes macro en estudios como el de Chen y otros (1986) apuntan hacia aquellos factores que inciden en las variaciones en la Tasa oficial de Descuento, entendiendo que esta es la que se emplea como expectativa de descuento futuro de los flujos de caja e, indirectamente, modifica el valor de estos. Las principales variables consideradas en el documento referido son:

- Diferencial entre el tipo de interés oficial a corto y largo plazo (prima de riesgo por plazo).
- Inflación esperada.
- Inflación realizada.
- Crecimiento de la producción industrial.
- Diferencial de rentabilidad entre los bonos de grandes y pequeñas empresas.

También se consideran, para ser finalmente desechadas, el crecimiento del consumo privado y los precios del petróleo.

Campbell (1993) profundiza en algunas consideraciones sobre estos factores considerados tradicionales. Su aportación se podría resumir en dos grandes ideas:

- El consumo privado agregado no es un buen representante del consumo tipo de los agentes que intervienen en los mercados de capitales. En el consumo agregado, la proporción de personas que no intervienen en los mercados de activos es muy alta por lo que, esta variable, registrará movimientos que nada tienen que ver con los mercados de capitales.
- En los mercados de capitales existen determinantes de compra-venta de dos naturalezas diferentes: por un lado estarán aquellos inversores que maximizan su curva de utilidad de una forma racional y, por otro, aquellos que producen operaciones más llevados por impulsos o "motivos psicológicos". Los primeros se verán movidos a acomodar sus carteras a las demandas y ofertas que producen estos "inversores impulsivos", rompiéndose los equilibrios lógicos teóricos del mercado.

En un interesante artículo de Schwert (1989) se hace un detallado repaso del papel explicativo de algunos acontecimientos macroeconómicos sobre la evolución de la volatilidad. Para Schwert, los principales factores que habría que considerar son los siguientes:

- Inflación y crecimiento de la masa monetaria.
- Nivel de actividad real; aunque el autor señala, en las conclusiones, que la volatilidad es mejor predictor de futuros vaivenes en el nivel de crecimiento que al revés.
- Efectos apalancamiento (dando lugar a una mayor duración de las volatilidades registradas). El autor recoge que la reducción de la rentabilidad de las acciones frente a los bonos o la emisión nueva de estos por encima de la emisión nueva de acciones da lugar a incrementos en la volatilidad.
- Variables de tipo comercio: laboralidad (días hábiles) y volumen de actividad comercial. Con sus valores lógicos, el autor demuestra que hay una relación directa



entre el número de días hábiles y el incremento de la volatilidad; ocurriendo lo mismo en cuanto a los días de mayor volumen de comercio relativo frente a los de menor volumen.

Desde el lado de la determinación de factores explicativos tipo micro, destaca la aportación y recopilación hecha por Fama y French (1993) y los mismos autores en (1996). Entre los tradicionalmente más utilizados se pueden contar:

- Valor de mercado de las acciones.
- Ratio precio-rentabilidad.
- Peso relativo de las empresas en el sector.
- Grado de exposición externa-interna.

Desde el punto de vista de la estimación estadística de los factores a incluir en el modelo, los métodos más empleados se refieren al análisis factorial, buscando aquella matriz de varianzas-covarianzas ortogonal que recoja toda la información de cada una de las variables incluidas y en forma decreciente; es decir, recogiendo el máximo de la varianza total en el primer factor y así sucesivamente hasta completar el total.

Con frecuencia, los factores explicativos de esta teoría se encuentran dentro de la propia matriz de varianzas-covarianzas entre todos los activos del mercado, lo que abre la puerta a modelos especialmente dedicados al estudio de la formación de esta y de sus posibilidades de cambio en el tiempo (modelos factor-ARCH de Engle (1987) o Engle y otros (1990). Este tipo de modelos se estudiará en profundidad más adelante.

### **Estimación y contraste de la APT**

Si definimos nuevamente el modelo como en la Ec. 176, tenemos:

$$\begin{aligned}
R_t &= a + Bf_{kt} + \varepsilon_t \\
E(\varepsilon_t) &= 0 \\
E(\varepsilon_t \varepsilon_t') &= \Sigma \\
E(f_{kt}) &= \mu_{fk} \\
E((f_{kt} - \mu_{fk})(f_{kt} - \mu_{fk})') &= \Omega_k \\
\text{cov}(f_{kt} \varepsilon_t') &= 0
\end{aligned}
\tag{Ec. 179}$$

donde,

$R_t$ : vector de (N x 1) de rentabilidades de los N activos de la cartera

B: matriz de (N x K) de sensibilidades

$f_{kt}$ : vector de (K x 1) de factores

$a$ : vector de términos constantes de (N x 1)

Los valores de los parámetros se estimarán por máxima verosimilitud, obteniéndose los siguientes resultados:

$$\begin{aligned}
\hat{a} &= \hat{\mu} - \hat{B}\hat{\mu}_{fk} \\
\hat{B} &= \left[ \sum_{t=1}^T (R_t - \hat{\mu})(f_{kt} - \hat{\mu}_{fk})' \right] x \left[ \sum_{t=1}^T (f_{kt} - \hat{\mu}_{fk})(f_{kt} - \hat{\mu}_{fk})' \right]^{-1} \\
\hat{\Sigma} &= \frac{1}{T} \sum_{t=1}^T (R_t - \hat{a} - \hat{B}f_{kt})(R_t - \hat{a} - \hat{B}f_{kt})' \\
\hat{\mu} &= \frac{1}{T} \sum_{t=1}^T R_t \quad \hat{\mu}_{fk} = \sum_{t=1}^T f_{kt}
\end{aligned}
\tag{Ec. 180}$$

De cara a contrastar la validez del modelo propuesto es válido el empleo de un ratio de máxima-verosimilitud similar al que se empleaba con la CAPM, de forma que podremos contrastar:

$$j = -(T - \frac{N}{2} - K - 1) [\log |\hat{\Sigma}| - \log |\hat{\Sigma}^*|] \tag{Ec. 181}$$

este estadístico se distribuye, cuando se cumple la hipótesis nula de valor nulo del término constante, como una  $\chi^2$  con grados de libertad iguales al número de restricciones impuestas.

El modelo restringido que se propone para la contrastación sería el siguiente:

$$R_t = i\gamma_0 + B\gamma_1 + Bf_{kt} + \varepsilon_t \quad \text{Ec. 182}$$

donde se está suponiendo que:

$$a = i\gamma_0 + B(\lambda_k + \mu_{fk}) \quad \text{Ec. 183}$$

Para el contraste de máxima-versedosimilitud, habrá N grados de libertad menos K necesarios para estimar los elementos de  $\lambda_k$  y 1 en la estimación de  $\gamma_0$ ; es decir N-K-1.

Los estimadores de este modelo restringido serían:

$$\begin{aligned} \hat{B}^* &= \left[ \sum_{t=1}^T (R_t - i\hat{\gamma}_0)(f_{kt} + \hat{\gamma}_1)' \right] X \left[ \sum_{t=1}^T (f_{kt} + \hat{\gamma}_1)(f_{kt} + \hat{\gamma}_1)' \right]^{-1} \\ \hat{\Sigma}^* &= \frac{1}{T} \sum_{t=1}^T \left[ (R_t - i\hat{\gamma}_0) - \hat{B}^*(f_{kt} + \hat{\gamma}_1) \right] \left[ (R_t - i\hat{\gamma}_0) - \hat{B}^*(f_{kt} + \hat{\gamma}_1) \right]' \\ \hat{\gamma} &= \left[ X' \hat{\Sigma}^{*-1} X \right]^{-1} \left[ X' \hat{\Sigma}^{*-1} (\hat{\mu} - \hat{B}^* \hat{\mu}_{fk}) \right] \\ X &= \begin{bmatrix} i\hat{B}^* \\ \end{bmatrix} \\ \gamma &= [\gamma_0 \gamma_1'] \end{aligned}$$

Aplicando la expresión de la Ec. 181 contrastaríamos el cumplimiento de la APT en la medida en que el estadístico creado estuviera incluido en la  $\chi^2_{N-K-1}$ .

### 6.3 Modelo APT con estructura FACTOR-ARCH de la matriz de varianzas-covarianzas entre los activos

Como ya se decía anteriormente, este tipo de modelos son susceptibles de una aplicación tipo ARCH para definir la prima de riesgo de cada activo en función de la matriz de varianzas-covarianzas temporal de todos los activos del mercado como factor explicativo.

Rescribiendo la formulación básica de la APT (Ec. 177) en lugar de para la rentabilidad bruta para el exceso de rentabilidad sobre el activo sin riesgo del mercado (prima de riesgo para cada activo) tenemos:

$$Z_t = \mu_t + \sum_{k=1}^K B f_{kt} + v_t \quad \text{Ec. 184}$$

donde  $\mu_t$  es la esperanza (media) condicional de  $Z_t$ , (vector  $N \times 1$  de primas de riesgo de cada uno de los activos). En esta expresión, los factores afectan a todas las primas de riesgo de los diferentes activos debidamente ponderados por sus  $b$ 's correspondientes, siendo  $v_t$  un vector de perturbaciones aleatorias. Podríamos escribir una matriz ( $N \times N$ ) de varianzas-covarianzas entre las primas de riesgo de los diferentes activos como:

$$H_t = \sum_{k=1}^K b_k b_k' \lambda_{kt} + \Omega \quad \text{Ec. 185}$$

donde,  $\Omega$  es una matriz semidefinida positiva de orden ( $N \times N$ ),  $\lambda_{kt}$  es un vector de ( $N \times 1$ ) de variables aleatorias y las  $b$ 's son linealmente independientes.

En la Ec. 185 se está haciendo un supuesto sobre la Ec. 184 que podría ser cualquiera de los dos casos siguientes:

1.  $B = b_k \lambda_{kt}^{1/2}$  y  $\text{var}_{t-1}(f_{kt}) = 1$ , situación en la que nos encontraríamos con un modelo en el que las covarianzas (betas) de los diferentes activos con un factor particular cambian proporcionalmente. Todos los activos responden de igual manera ante el cambio en un determinado factor.
2.  $B = b_k$  y  $\text{var}_{t-1}(f_{kt}) = \lambda_{kt}$ , donde la respuesta de cada prima de riesgo ante la variación de un factor es distinta.

La matriz  $H_t$  goza de una serie de propiedades de enorme utilidad para su utilización:

- Es semidefinida positiva.
- Siempre se pueden construir carteras de valores con matriz de varianzas-covarianzas condicional constante.
- Siempre se puede reemplazar el  $\lambda_{kt}$  de la Ec. 185 transformando la expresión del siguiente modo:

$$H_t = \sum_{k=1}^K b_k b_k' h_{kt} + \Omega^* \quad \text{Ec. 186}$$

donde  $h_{kt}$  son las varianzas condicionales de las carteras de valores de N activos y  $\Omega^*$  es una matriz de (N x N) constante.

Si construimos una cartera de valores  $P_{kt} = \alpha_k' Z_t$ , donde  $\alpha_k$  es el vector de pesos de cada activo en la cartera, de forma que  $\alpha_k' b_k = 1$ , (lo que siendo el vector de betas ortogonal supone que el valor de alfa es:  $\alpha_k = b_k / b_k' b_k$ ), la varianza condicional de la prima de riesgo de esta cartera se puede escribir como:

$$h_{kt} = \alpha_k' H_t \alpha_k = \lambda_{kt} + \alpha_k' \Omega \alpha_k = \lambda_{kt} + s_k \quad \text{Ec. 187}$$

En definitiva, podemos escribir la matriz de varianzas-covarianzas como:

$$H_t = \sum_{k=1}^K b_k b_k' h_{kt} + \Omega^* \quad \text{Ec. 188}$$

donde,  $\Omega^* = \Omega - \sum_{k=1}^K b_k b_k' s_k$ .

- Se pueden obtener valores a futuro de la matriz de varianzas-covarianzas a partir de la predicción multiperíodo del vector  $h_{kt}$ . Su expresión sería la siguiente:

$$E_t(H_{t+\tau}) = \sum_{k=1}^K b_k b_k' E_t(h_{kt+\tau}) + \Omega^* \quad \text{Ec. 189}$$

De cara a la especificación econométrica del modelo y las relaciones causales hasta aquí planteadas es necesario llevar a cabo una definición concreta del vector  $h_{kt}$ . Para ello, Engle y otros (1990) proponen dos supuestos simplificadores, ya que, considerar todos los factores o carteras posibles daría lugar a un sistema imposible de estimar. Las propuestas serían las siguientes:

1. Considerar que el conjunto de las K carteras o factores tienen una representación univariante. Esto es lo mismo que asumir que la prima de riesgo de cada cartera (factor) "k" está bien definida por una función univariante que puede ser, por ejemplo, un esquema GARCH (1,1) en Media del tipo:

$$\begin{aligned} P_{kt} &= c_k + \gamma_k h_{kt} + u_{kt} \\ u_{kt} | \psi_{t-1} &\Rightarrow N(0; h_{kt}) \\ h_{kt} &= w_k + \alpha_k u_{kt-1}^2 + \beta_k h_{kt-1} \end{aligned} \quad \text{Ec. 190}$$

dado que  $u_{kt} = \alpha_k' \varepsilon_t$ , y que  $h_{kt}$  y  $\lambda_{kt}$  están relacionados por  $h_{kt} = \lambda_{kt} + s_k$  (tal y como se veía en la Ec. 187), la matriz de varianzas-covarianzas se puede escribir como:

$$\begin{aligned} H_t &= C^* + \sum_{k=1}^K [\alpha_k b_k b_k' (\alpha_k' \varepsilon_{t-1}^2)^2 + \beta_k b_k b_k' (\alpha_k' H_{t-1} \alpha_k)] \\ C^* &= \sum b_k b_k' w_k + \Omega^* \end{aligned} \quad \text{Ec. 191}$$

En Baba y otros (1987) se demuestra que  $h_{kt}$  se puede estimar consistentemente por máxima-verosimilitud. Las b's se pueden calcular usando estos valores, obteniéndose estimadores consistentes, pero no eficientes.

2. La segunda alternativa es considerar un esquema multivariante para la formación del factor "k", que será calculado siguiendo un esquema recursivo. El sistema propuesto admite la posibilidad de que el factor en "t" dependa de la información existente de su propio pasado y del comportamiento observado de los primeras k-1 carteras consideradas. La forma de estimar este caso es algo más compleja y se detalla, por ejemplo, en Engle (1987).



## **7. CONTRASTES DE ESTRUCTURA Y BONDAD DE AJUSTE EN LOS MODELOS ARCH**



## 7.1 Introducción

Siguiendo la metodología de Hendry, la contrastación de un modelo econométrico pasa por dos fases:

- una primera en la que se contrasta "mala especificación"; es decir, la aceptación de la hipótesis nula no lleva a la aceptación de la hipótesis alternativa,
- y una segunda de "especificación" propiamente dicha, en la que la hipótesis alternativa es "constructiva": rechazar la hipótesis nula supone aceptar la alternativa

Pérez y Murillo (1997) clasifican las distintas medidas de contrastación del siguiente modo:

Mala especificación	No linealidad Heterocedasticidad	
	Sobre la curva de las noticias	Sesgo de signo Sesgo signo positivo Sesgo signo negativo Conjunto
	Capacidad predictiva	
	Otros (varianza, normalidad, asimetría, ...)	
	Comparación entre modelos	Modelos anidados Modelos no anidados
Especificación	Estabilidad varianzas	

Siguiendo su mismo esquema de organización aunque incluyendo algunos estadísticos más, hacemos un breve recorrido por las diferentes fases en la contrastación del modelo ARCH.

## 7.2 Contrastes de mala especificación

### Contrastes de no linealidad

Désde el artículo seminal de Engle, tantas veces citado, ya se desprende que la determinación de la no linealidad en el comportamiento del error es la clave de los

modelos ARCH y, obviamente, debe ser la primera comprobación que nos incite a su empleo.

Continuamente estamos hablando de la posibilidad de la existencia de una correlación entre los valores al cuadrado del error en distintos momentos del tiempo como determinantes de una varianza condicional heterocedástica autorregresiva. Siguiendo la metodología de Box-Jenkins, es inmediata la comprobación de este hecho a partir de las funciones de autocorrelación total y parcial cuyos casos concretos ya han sido comentadas y definidas en diversas secciones de esta tesis con motivo de la exposición de cada tipología ARCH diferente.

Estimada la función de autocorrelación total de un proceso aleatorio  $X$  (que puede perfectamente ser la perturbación al cuadrado) como:

$$r_k = \frac{\sum_{t=k+1}^N (x_t - \bar{x})(x_{t-k} - \bar{x}_{t-k})}{NS^2}, k = 1, 2, \dots \quad \text{Ec. 192}$$

Existen multitud de criterios para poder definir si un proceso estocástico es un ruido blanco o no. De entre ellos, los más extendidos (citados por orden cronológico de aparición) son:

### Intervalo de confianza

Para distribuciones fijas con varianza finita ( $\rho_k=0$  /  $k \neq 0$ ), la varianza de un valor de la función de autocorrelación total  $r_k$  viene dada aproximadamente por el valor  $N^{-1}$ . Si, además, el recorrido muestral es lo suficientemente largo (si  $N$  es grande), será aproximadamente una normal con lo que un valor de  $r_k$  por encima de  $2 N^{-1/2}$  puede considerarse como significativamente distinto de cero.

De un modo más general, la varianza para un número de parámetros superior a "q" se podría especificar como:

$$\text{Var}(r_k) = N^{-1} \cdot (1 + 2\rho_1^2 + \dots + 2\rho_q^2). \quad \text{Ec. 193}$$

El valor muestral fuera de las dos bandas que definen  $\pm 2$  veces la desviación típica (para el caso de una normal) mostrarían que no estamos ante un ruido blanco (o lo que es lo mismo, que existe algún valor de la función de autocorrelación total significativamente distinto de cero, ya que excede los límites en los que se encontraría la normal (0,1) según la cual se ha de distribuir un ruido blanco).

### Cociente de Von Neumann:

Es un test construido para muestras muy grandes, bajo el supuesto de que, en ese caso, la siguiente expresión se distribuye como una normal

$$\frac{\delta^2}{S^2} = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2 / (n-1)}{\sum (e_i - \bar{e})^2 / n} \quad \text{Ec. 194}$$

Cuando se ha realizado la estimación de determinada forma, la media de los errores es nula y, con un "n" suficientemente grande, se puede suponer que la distribución anterior se distribuye como una normal con media y varianza como siguen:

$$E\left(\frac{\delta^2}{S^2}\right) = \frac{2n}{n-1} \quad \text{Var}\left(\frac{\delta^2}{S^2}\right) = \frac{4n^2(n-2)}{(n+1)(n-1)^3} \quad \text{Ec. 195}$$

Las fórmulas de media y varianza son sólo ciertas en el caso en que los errores se distribuyan de forma independiente, con lo cual el cociente tiene una validez muy limitada.

### D y H de Durbin-Watson:

Este estadístico ya no fundamenta su base en el límite, con lo que es válido para muestras más pequeñas (valores de "n" menores).

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad \text{Ec. 196}$$

A partir de una transformación muy simple, se puede llegar a definir esta "d" de Durbin-Watson como una función casi equivalente del coeficiente de autocorrelación total:

$$d \approx 2(1 - \hat{\rho}) \quad \text{donde} \quad \hat{\rho} = \frac{\sum_{i=2}^n e_i e_{i-1}}{\sum_{i=1}^n e_i^2} \quad \text{que se demuestra es el estimador insesgado del}$$

coeficiente de autocorrelación total.

En el caso en que la variable endógena retardada está entre las variables explicativas del modelo, el estadístico anterior no se puede emplear, ya que, parte del autocorrelación existente entre las perturbaciones aleatorias correspondientes a distintas observaciones, queda "camuflada" por la inclusión de  $y_{i-1} = f(u_{i-1})$ , en la que  $u_i = f(u_{i-1})$ .

En este caso, la "d" expresada anteriormente hablaría de ausencia de autocorrelación aún cuando ésta existiera, por lo cual se define una nueva expresión que sería:

$$h = \hat{\rho} \sqrt{\frac{n}{1 - nS^2(\hat{\beta}_{y_{i-1}})}} \quad \text{Ec. 197}$$

En la cual el coeficiente de autocorrelación se aumenta con una función que incluye la varianza del coeficiente de la variable endógena desplazada que funciona como explicativa, así como con el número de observaciones del modelo. Obviamente, este contraste es válido siempre y cuando el denominador de la expresión sea mayor que cero, por lo que necesariamente la varianza del parámetro de la endógena retardada debe ser inferior a  $N^{-1}$ . Si esta condición no se diera, los autores recomiendan realizar una regresión de la perturbación aleatoria en función de todas las demás explicativas (incluida en su caso la endógena desplazada) y de ella misma desplazada ( $U_{i-1}$ ) y determinar la presencia o no de autocorrelación en función de la significatividad del parámetro de esta última explicativa, juzgada esta por el valor de la t-student.

Ambas expresiones presentadas por Durbin y Watson se distribuyen asintóticamente como una  $\chi^2$  con "k" (número de variables explicativas) grados de libertad. Aún así, dado que el valor estimado de la perturbación aleatoria es una combinación lineal de la propia perturbación y de las variables explicativas del modelo, no se pueden obtener valores exactos para la exclusión o inclusión del supuesto de autocorrelación. Sólo se pueden construir unos valores críticos (superiores e inferiores) que determinen, en función del número de datos y del número de variables explicativas, unos intervalos de total seguridad de existencia o no y otros de zona de incertidumbre.

Este estadístico y el anterior sólo son válidos para cuando se supone que la autocorrelación del término de la perturbación aleatoria se puede explicitar como un modelo autorregresivo de orden uno.

#### Q estadística de Box Pierce :

Con esta expresión se estaría obteniendo un valor crítico único del valor global de cada autocorrelación calculada en función de una distribución tabulada. En la expresión se premian los modelos con mayor número de datos (está multiplicada por N) y se puede demostrar que la función estadística del valor calculado se distribuye como una chi-cuadrado con "k" grados de libertad

$$Q(k) = N \cdot \sum_{i=1}^k r_i^2 \Rightarrow \chi_k^2 \quad \text{Ec. 198}$$

Con este criterio, valores de la "Q" experimental calculada según esta expresión superiores al valor tabulado de la chi-cuadrado nos llevarían a rechazar la existencia de un ruido blanco (los valores de la función de autocorrelación serían, globalmente, diferentes de cero).

#### Q estadística de L-Jung-Box :

A partir de una serie de simulaciones sobre la Q de Box-Pierce antes citada, se comprobó que este estadístico podría estar infravalorando el nivel de significación a partir del cual se

debe rechazar la existencia de un ruido blanco. Por ello se propuso la corrección a la fórmula anterior que ahora se cita y que intenta paliar este problema de infravaloración.

$$Q^*(k) = N(N+2) \cdot \sum_{i=1}^k (N-i)^{-1} \cdot r_i^2 \Rightarrow \chi_k^2 \quad \text{Ec. 199}$$

Obviamente, la interpretación del estadístico calculado sigue las mismas reglas que las expresadas para el caso de la Q de Box-Pierce.

#### Test de correlación serial del Multiplicador de Lagrange. (Breusch-Godfrey):

Estos autores proponen la estimación del modelo, mediante mínimos cuadrados ordinarios, dos veces, una para obtener los errores y otra segunda del modelo inicial más una especificación accesoria en función de los errores retardados obtenidos:

$$1. y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \beta_{k+1} y_{t-1} + \beta_{k+2} y_{t-2} + \dots + U_t \text{ donde } e_t = y_t - \hat{y}_{t-1}, \text{ y}$$

$$2. y_t = \beta_1 + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + \beta_{k+1} y_{t-1} + \beta_{k+2} y_{t-2} + \dots + \varphi_1 e_{t-1} + \varphi_2 e_{t-2} + \dots + \varphi_p e_{t-p} + U_t$$

Los autores demuestran que la expresión  $T \cdot R^2$  (T: tamaño de la muestra y  $R^2$ : coeficiente de correlación lineal de la regresión) se distribuye asintóticamente como una  $\chi^2$  con "p" grados de libertad. Se rechazará la hipótesis nula de no existencia de autocorrelación si esta expresión es superior al valor de tablas.

Este contraste es equivalente a la Q de Box-Pierce cuando el vector de variables explicativas contiene únicamente variables exógenas.

#### Criterio informativo de Akaike:

Con este criterio se penaliza la existencia de un mayor número de variables explicativas aunque el valor de esta penalización puede verse compensado por la ganancia en bondad

de ajuste medida por la varianza del error obtenido en la experimentación. Dicho de otra forma, se aceptará un modelo con más términos frente a otro siempre que la ganancia por reducción en la variación del error compense "el daño" por disminuir los grados de libertad del modelo.

$$AIC = \frac{2k}{N} + \log\left(\frac{e'e}{N}\right) \quad \text{Ec. 200}$$

De un modo aplicado, el criterio nos servirá para rechazar un modelo con más variables explicativas que otro cuando la inclusión de más variables no haga que el error disminuya de forma considerable. En definitiva, siempre será preferible un modelo frente a otro cuando de este obtengamos un valor AIC más pequeño.

#### **Modelo de selección de portfolio de Poskitt y Tremayne:**

Estos dos autores introducen una expresión para determinar cuál debe ser el modelo seleccionado en función de los valores obtenidos por el criterio de Akaike. Sugieren la siguiente expresión:

$$\mathfrak{R} = \exp\left(-\frac{1}{2}T(AIC(p_1, q_1) - AIC(p, q))\right) \quad \text{Ec. 201}$$

Los autores de esta expresión refieren que un valor de esta expresión menor que  $\sqrt{10}$  sería significativo de que la combinación de parámetros  $(p_1, q_1)$  son parámetros igualmente aceptables que la combinación  $(p, q)$ .

#### **Criterio de información de Schwarz:**

Se trata de una variante del criterio de Akaike, en esta ocasión imponiendo una mayor penalización a un modelo frente a otro por la introducción de una variable explicativa más.

$$BIC = \frac{k \log(N)}{N} + \log\left(\frac{e'e}{N}\right) \quad \text{Ec. 202}$$

Obviamente, la interpretación es la misma que la ya mencionada para el criterio de Akaike.

Tanto el criterio de información de Akaike como el de Schwarz no se diseñaron específicamente para dirimir la cuestión de si estamos o no ante un ruido blanco, sino para medir de alguna forma el posible problema de sobreparametrización de un modelo econométrico, dicho de otra forma, para saber cuál es el largo óptimo del número de retardos que han de introducirse en la especificación de un proceso. Aún así, su utilidad es obvia por cuanto nos dirá si la inclusión de una explicativa más en el modelo compensa el perjuicio causado por la reducción de los grados de libertad del modelo en función de la reducción en la desviación típica del error cometido en el ajuste.

### Contraste de heterocedasticidad

A partir de la propia definición de un modelo GARCH(p,q) está implícita la contrastación sobre la heterocedasticidad, por cuanto la nulidad de los parámetros de la parte autorregresiva en la especificación de la varianza condicional servirían para determinar este punto.

Esta situación da lugar al contraste de heterocedasticidad conocido con el nombre de LM ARCH, que podría expresarse formalmente como:

Teniendo un modelo GARCH(p,q) del tipo:

$$\begin{aligned} h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2} + \dots + \beta_p h_{t-p} = \\ &= Z_1' \theta_1 + Z_2' \theta_2 \end{aligned}$$

correspondiendo  $\theta_2$  a la parte de media móvil (parámetros) del modelo anterior, el contraste de heterocedasticidad sería tan sencillo como comprobar si este vector de parámetros es nulo o no, para lo cual se propone el test máxima verosimilitud en su versión abreviada por Breusch y Pagan que ya se comentaba en la sección anterior.



### La curva de impacto de las innovaciones previas (*News Impact Curve*)

Engle y Ng (1993) propusieron una medida del grado de ajuste entre dos o más modelos tipo ARCH a partir de la llamada Curva de Impacto de las Innovaciones (*News Impact Curve*) que definieron del siguiente modo:

*... Manteniendo constante la información referida a  $t-2$  y la previa a este punto, se puede examinar la relación implícita entre  $\varepsilon_{t-1}$  y  $h_t$ . Llamaremos Curva de Impacto de las Innovaciones a la curva con todos las varianzas condicionales retardadas frente al nivel de la varianza marginal del proceso, lo que sirve para relacionar los "shocks" del pasado (innovaciones) con la volatilidad estimada*

Las dos virtudes principales imputables a esta curva se refieren a:

- servir para poner de manifiesto las diferencias entre los distintos modelos ARCH de cara a la predicción de la volatilidad, y
- contrastar la bondad del modelo a partir del grado de ajuste que la curva de impacto de las innovaciones tiene sobre los datos reales.

Para construir la curva a la que nos estamos refiriendo, se siguen los siguientes pasos:

1. Se estima el modelo bajo cualquiera de las propuestas ya comentadas (ARCH, GARCH, GJR-ARCH, EGARCH, AGARCH, etc.). Así se obtiene el valor de los parámetros.
2. Se obtiene la curva de Impacto de las Innovaciones calculando la varianza condicional según la especificación que se le haya dado a ésta. En el cuadro siguiente, se expresa como se obtiene ésta para las principales modificaciones de los modelos ARCH, pudiendo obtenerse cualquier otra a partir de la omisión de alguno de los términos en las aquí referidos.

Cuadro : Modelización y cálculo de la Curva de Impacto de las Innovaciones<sup>17</sup>

	Especificación Varianza	Curva de Impacto de las noticias
GARCH(1,1)	$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$	$h_t = A + \hat{\alpha}_1 \varepsilon_{t-1}^2$ $A = \hat{\alpha}_0 + \hat{\beta}_1 \sigma^2$
EGARCH Logarítmico	$\log(h_t) = \alpha_0 + \beta_1 \log(h_{t-1}) + \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \alpha_1 \left[ \frac{ \varepsilon_{t-1} }{\sqrt{h_{t-1}}} - \sqrt{2/\pi} \right]$	$h_t = \begin{cases} A \exp \left[ \frac{(\gamma + \hat{\alpha}_1)}{\sigma} \varepsilon_{t-1} \right] & \text{si } \varepsilon_{t-1} > 0 \\ A \exp \left[ \frac{(\gamma - \hat{\alpha}_1)}{\sigma} \varepsilon_{t-1} \right] & \text{si } \varepsilon_{t-1} < 0 \end{cases}$ $A = \sigma^{\beta_1} \exp(\hat{\alpha}_0 - \hat{\alpha}_1 \sqrt{2/\pi})$
GJR-ARCH	$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} + \gamma d_{t-1} \varepsilon_{t-1}^2$ $d_{t-1} = 1 \text{ si } \varepsilon_{t-1} < 0 \text{ y } 0 \text{ resto}$	$h_t = \begin{cases} A + \hat{\alpha}_1 \varepsilon_{t-1}^2 & \text{si } \varepsilon_{t-1} > 0 \\ A + (\hat{\alpha}_1 + \gamma) \varepsilon_{t-1}^2 & \text{resto} \end{cases}$ $A = \hat{\alpha}_0 + \hat{\beta}_1 \sigma^2$
AGARCH	$h_t = \alpha_0 + \alpha_1 (\varepsilon_{t-1} + \gamma)^2 + \beta_1 h_{t-1}$	$h_t = A + \hat{\alpha}_1 (\hat{\gamma} + \varepsilon_{t-1})^2$ $A = \hat{\alpha}_0 + \hat{\beta}_1 \sigma^2$

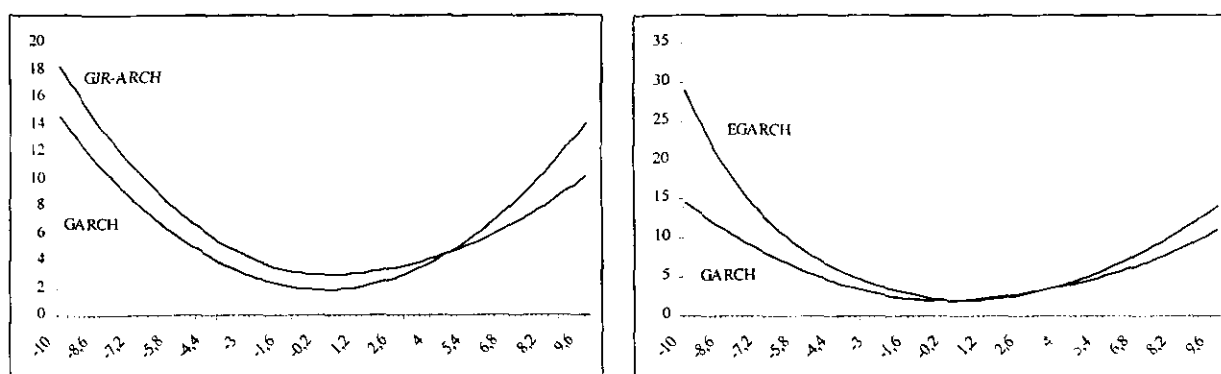
Realizando un gráfico de las diferentes curvas correspondientes a los modelos más comúnmente utilizados (GARCH, EGARCH y GJR-GARCH) - ver secciones correspondientes - se pueden observar las siguientes características:

- Como es obvio y ya se ha señalado en varias ocasiones, el proceso GARCH es simétrico, no habiendo ninguna diferencia en la modelización de la varianza a partir de la constatación de que el diferente signo de  $\varepsilon_{t-1}$  puede producir diferentes efectos. Los otros dos modelos propuestos (EGARCH y GJR-ARCH) sí diferencian entre ambas situaciones (signo positivo o negativo).
- En los modelos EGARCH, grandes valores de la innovación (perturbación) en el período anterior derivan en efectos sobre la volatilidad mayores que en el modelo GARCH.

<sup>17</sup> En este cuadro,  $\alpha_i, \beta_i, \gamma$  son parámetros constantes y estimables; mientras que  $\sigma^2$  es la varianza marginal del proceso y  $h_t$  es la varianza condicional que se modeliza.

- Los tres modelos propuestos están centrados en  $\varepsilon_{t-1} = 0$ , aunque, tanto el modelo GJR-ARCH como el EGARCH, presentan distintas pendientes para los valores positivos y para los negativos (más pronunciada esta última), siendo mayor la diferencia entre ambas en el modelo GJR-ARCH.

**Ilustración 1: Curva de Impacto de las Innovaciones para distintos modelos ARCH**



Ambos gráficos están elaborados sobre los datos diarios del IBEX-35 de Madrid entre enero de 1992 y junio de 1999. El modelo GARCH es de orden (1,1) y el modelo GJR-ARCH no incluye ninguna variable explicativa accesoria sobre la varianza. La forma de cálculo de las curvas se recoge en el cuadro previo, aunque, de cara a la observación del componente teórico, se ha asumido como valor conocido de la varianza condicional en el período  $t-2$  el valor de la mediana de la varianza estimada según las distintas propuestas (se asume, entonces, que los procesos son Gausianos)

Dadas las diferencias manifiestas entre la modelización simétrica de la varianza condicional o volatilidad (representada por el proceso GARCH) y las asimétricas (EGARCH y GJR-ARCH), es necesario determinar en qué medida la realidad cumple la hipótesis de impacto diferente ante innovaciones negativas y positivas. Engle y Ng propusieron un triple contraste de posibles sesgos diferentes:

- Contraste sobre el sesgo del signo (o conjunto), centrado en estudiar si hay impactos en el modelo que no considera variaciones por el signo de las innovaciones en el periodo anterior y, por tanto, éste está mal especificado

- Contraste sobre el sesgo negativo, en el cual se intenta comprobar el efecto distinto en la volatilidad resultante producido por la diferencia entre valores negativos pasados grandes y pequeños de la perturbación aleatoria.
- Contraste sobre el sesgo positivo, igual que el caso anterior, pero ahora para las innovaciones previas positivas.

En definitiva, los tres contrastes propuestos introducirán en la especificación inicial (simétrica) variables explicativas que recojan efectos que, de ser significativos, nos determinarían una mala especificación del modelo inicial.

Contraste	Objetivo	Variable introducida
Sesgo en el signo	existencia o no de asimetría	$S_{t-1}^- = \begin{cases} 1 & \text{si } \varepsilon_{t-1} < 0 \\ 0 & \text{resto casos} \end{cases}$
Sesgo negativo	efectos de "apalancamiento" (leverage) en valores negativos	$S_{t-1}^- \varepsilon_{t-1}$
Sesgo positivo	efectos de "apalancamiento" en valores positivos	$S_{t-1}^+ = 1 - S_{t-1}^-$

Este tipo de contrastes es fácilmente generalizable a especificaciones de las variables ficticias comentadas que toman valor uno si el valor de la innovación en el período precedente excede de un tanto por ciento dado de la estructura de sus percentiles. El tipo de contraste propuesto obviamente no es tan sólo apto para modelos tipo ARCH y puede establecerse de forma separada o conjunta.

Formalmente, el contraste se expresaría del siguiente modo:

Sea la especificación de la varianza condicional dada por una función del siguiente tipo:

$$\log(h_t) = \log(h_0 + (\delta_0' z_{0t})) + \delta_\alpha' z_{\alpha t} \quad \text{Ec. 203}$$

donde,

$h_t$ : varianza condicional ó  $E(\varepsilon_t^2)$ .

$h_{0t}(\delta_0' z_{0t})$ : modelo de la volatilidad bajo la hipótesis nula con  $\delta_{0t}$  un vector de  $(k \times 1)$  parámetros y  $z_{0t}$  un vector de  $(k \times 1)$  variables explicativas.

$\delta_{at}$ : vector de  $(m \times 1)$  parámetros de variables explicativas no incluidas en la especificación inicial y  $z_{at}$  las variables no incluidas.

La hipótesis a contrastar será obviamente la de nulidad de los parámetros adicionales:  $h_0(\delta_a)$ . Si se considera la perturbación del modelo ARCH normalizada, es decir  $v_t = \varepsilon_t / \sqrt{h_t}$ , se puede definir una regresión auxiliar del tipo de la enunciada por Engle en su artículo seminal como:

$$v_t^2 = \delta_0' z_{0t}^* + \delta_a' z_{at}^* + u_t \quad \text{Ec. 204}$$

donde  $z_{0t}^* = h_{0t}^{-1} \partial h_t / \partial \delta_0$  y  $z_{at}^* = h_{0t}^{-1} \partial h_t / \partial \delta_a$ . Se evalúan (por máxima-verosimilitud) que  $\delta_a = 0$  y se obtienen los valores estimados de  $\delta_0$  bajo la hipótesis nula. Empleando la habitual expresión de Breusch y Pagan (1978), se puede calcular el estadístico:

$$\zeta_{LM} = T \cdot R^2 \Rightarrow \chi_m^2 \quad \text{Ec. 205}$$

donde "m" es el número de variables explicativas accesorias cuya existencia en el modelo se desea contrastar,  $R^2$  es el coeficiente de correlación lineal de la Ec. 204 y T es el tamaño muestral.

Dado que las relaciones implícitas en un modelo tipo ARCH son no lineales, la ortogonalidad entre  $v_t$  y  $z_{0t}^*$  no está garantizada, razón por la que este contraste se suele llevar a cabo con el siguiente proceso:

1. Regresión de  $\varepsilon_t^2$  en función de  $z_{0t}$  solamente.
2. Usar los residuos de esta regresión (con lo que queda garantizada la ortogonalidad entre  $v_t$  y  $z_{0t}^*$ ).

### Contrastes sobre la capacidad predictiva del modelo

Para determinar diferencias significativas entre los diversos modelos ARCH posibles, algunos autores recogen distintas funciones de pérdida que evalúan la gravedad del error cometido. De entre ellos, los más comunes son los que desarrollaran Hamilton y Susmel (1994) y que se resumen en las siguientes cuatro propuestas:

	<b>Directo</b>	<b>Valores Absolutos</b>
<b>Errores</b>	$MSE = \frac{\sum_{i=1}^T (\varepsilon_i^2 - h_i)^2}{T}$	$MAE = \frac{\sum_{i=1}^T  \varepsilon_i^2 - h_i }{T}$
<b>Logaritmo de errores</b>	$LMSE = \frac{\sum_{i=1}^T (Ln(\varepsilon_i^2) - Ln(h_i))^2}{T}$	$LMAE = \frac{\sum_{i=1}^T  Ln(\varepsilon_i^2) - Ln(h_i) }{T}$

Donde  $h_i$  es la varianza condicional de  $\varepsilon_i^2$

La buena capacidad predictiva de un modelo obviamente se centrará en que el valor de estos estadísticos se encuentre próximo a cero. Los autores descritos, advierten sobre la imposibilidad de cálculo del error estándar medio (MSE) cuando estamos ante un IGARCH, ya que este se centra en el cálculo de los cuartos momentos que, en dicho proceso, no existen.

### Otros contrastes de mala especificación

En diversas situaciones y aportaciones metodológicas de la familia de los ARCH que se han venido explicando en capítulos anteriores, se ha puesto especial importancia en la forma funcional de distribución probabilística de los procesos que se iban a modelizar. Este punto es fundamental por se la clave de la estimación máximo verosímil que se realizará en cualquier aplicación de este tipo.

Como ya se ha repetido con frecuencia, las series financieras, a las que se les ha aplicado con mayor éxito este tipo de modelos, presentan con frecuencia una distribución más similar a la de la t-student que a la de la normal o a una distribución tipo GED como

la analizada por Nelson (1991). Es por ello que la comprobación de los estadísticos básicos sobre el comportamiento normal o no de la serie a modelizar es un paso obligado.

### **Coeficiente de asimetría**

Se define como coeficiente de asimetría (A) el momento de orden tres respecto a la media dividido por la desviación típica al cubo del proceso. Formalmente:

$$\begin{aligned} k &= \frac{m_3}{\sigma^3} = \frac{1}{\sigma^3} E(\xi - E(\xi))^3 = \frac{1}{\sigma^3} \int_{-\infty}^{\infty} (\xi - E(\xi))^3 dF(\xi) = \\ &= \frac{1}{\sigma^3} \int_{-\infty}^{\infty} (\xi - E(\xi))^3 f(\xi) d\xi \end{aligned} \quad \text{Ec. 206}$$

Dicho coeficiente de asimetría, calculado para la función de una normal, tiene un valor igual a cero (la función normal es perfectamente simétrica).

### **Coeficiente de curtosis o apuntamiento**

Se define como coeficiente de curtosis (k) el momento de orden cuatro respecto a la media dividido por la varianza del proceso. Formalmente:

$$\begin{aligned} k &= \frac{m_4}{\sigma^4} = \frac{1}{\sigma^4} E(\xi - E(\xi))^4 = \frac{1}{\sigma^4} \int_{-\infty}^{\infty} (\xi - E(\xi))^4 dF(\xi) = \\ &= \frac{1}{\sigma^4} \int_{-\infty}^{\infty} (\xi - E(\xi))^4 f(\xi) d\xi \end{aligned} \quad \text{Ec. 207}$$

Dicho coeficiente de apuntamiento o curtosis, calculado para la función de una normal, tiene un valor igual a tres.

### **Contraste de normalidad de Jarque Bera (1981)**

Ambos autores prueban que el estadístico experimental:

$$JB = N \left( \frac{(\text{asimetría})^2}{6} + \frac{(\text{curtosis} - 3)^2}{24} \right) \rightarrow \chi^2_2 \quad \text{Ec. 208}$$

se distribuye como una  $\chi^2$  con dos grados de libertad cuando la serie sobre la que se aplica se distribuye como una normal.

### 7.3 Contrastes de especificación

#### Contrastes de ratio de verosimilitud, Multiplicador de Lagrange y Test de Wald

En principio, estos contrastes se consideran "constructivos" ya que, tal y como se citaba anteriormente, el rechazo de la hipótesis nula que se plantea supone la aceptación de la alternativa.

En la literatura, son frecuentes el empleo de las tres variantes básicas de contrastación de un modelo restringido sobre uno aumentado a partir del grado de "máximo" de la función de verosimilitud del modelo. Dicho de otra forma, los tres contrastes citados responden a la misma idea básica: si con el modelo restringido, y una vez estimados los parámetros correspondientes, estamos ante un máximo de la función de verosimilitud, la inclusión de nuevas variables explicativas no debe producir un crecimiento significativo de esta función.

Los contrastes más habituales son la razón de verosimilitud, el multiplicador de Lagrange y el test de Wald.

El hecho de emplear distintas propuestas sobre la misma contrastación se debe al hecho matemático de que una formulación que tan sólo responde a la primera derivada de una función puede estar minusvalorando o sobrevalorando los resultados por el grado de concavidad o convexidad matemáticas de la función que estemos empleando.

En el contraste del ratio de verosimilitud se plantea que si el modelo restringido (con menos parámetros) es suficiente, la ganancia en la maximización del logaritmo de verosimilitud al incluir más parámetros (más variables explicativas) no es significativa.



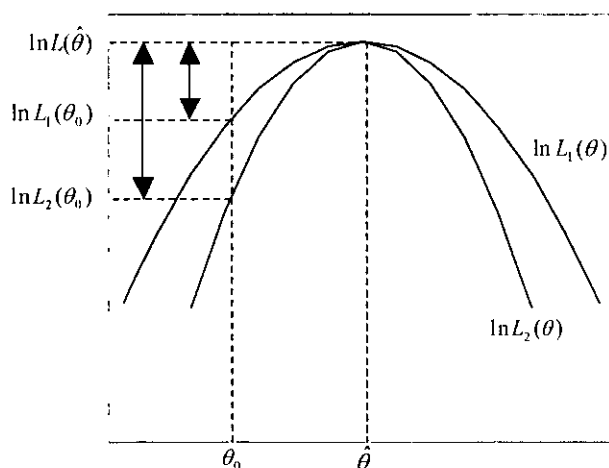
Habitualmente se emplea la siguiente expresión para su determinación:

$$RV = -2(\ln(H_0) - \ln(H_a)) \Rightarrow \chi_p^2 \text{ Ec. 209}$$

donde  $H_0$  es el valor del logaritmo de máxima verosimilitud estimado con el modelo restringido y  $H_a$  con el modelo ampliado o alternativo. Teóricamente, y bajo el supuesto contemplado en la hipótesis nula, el estadístico experimental se debe distribuir como una  $\chi_p^2$ , donde "p" es el número de variables explicativas que se incorporan en el modelo ampliado.

Tal y como ya se ha comentado, el empleo del ratio de verosimilitud da lugar a un sesgo en los resultados en función del grado de convexidad que presente la función, hecho que con este ratio no se está teniendo en cuenta.

Recogiendo la misma exposición de Buse (1982)<sup>18</sup>, en el gráfico siguiente se representan los logaritmos de verosimilitud normal de dos series, una con mayor convexidad que la otra:



<sup>18</sup> BUSE, A. (1982): *The Likelihood Ratio, Wald and Lagrange Multiplier Tests: An Expository Note*. The American Statistician, 36. Part I. 1982.

Sobre las relaciones de estos tres contrastes se puede encontrar un análisis detallado en OTERO, JM (1993): *Econometría y Series Temporales*. Editorial AC.

donde:

$\ln L(\hat{\theta})$ : valor del logaritmo de verosimilitud de la función uno para el parámetro  $\hat{\theta}$  (el máximo verosímil).

$\ln L_1(\theta_0)$ : valor del logaritmo de verosimilitud de la función uno para el parámetro  $\theta_0$

$\ln L_2(\theta_0)$ : valor del logaritmo de verosimilitud de la función uno para el parámetro  $\theta_0$

Es sencillo comprobar que el ratio de verosimilitud para el parámetro  $\theta_0$  es tanto mayor cuanto mayor es el grado de convexidad de la serie. Es decir, cuanto más convexo es el logaritmo, la posibilidad de rechazar la hipótesis nula en la ecuación siguiente es tanto mayor.

$$RV = -2(\ln(\hat{\theta}) - \ln(\theta)) \Rightarrow \chi_p^2 \quad \text{Ec. 210}$$

De cara a solventar este sesgo hacia el rechazo de la hipótesis nula cuando se comparan distintos logaritmos, se emplean tradicionalmente las dos propuestas de Wald y del Multiplicador de Lagrange.

En el test de Wald se emplea como medida de la significatividad la diferencia entre cada par de parámetros distintos matizada por la matriz de información (matriz de segundas derivadas) del siguiente modo:

$$W = r(\hat{\theta})' [R(\hat{\theta}) I(\hat{\theta})^{-1} R(\hat{\theta})]^{-1} r(\hat{\theta}) \rightarrow \chi_q^2 \quad \text{Ec. 211}$$

donde,

$r(\theta)$  = restricciones paramétricas planteadas

$R(\theta) = \frac{\partial r(\theta)}{\partial \theta}$ , matriz de derivadas parciales de (q x k)

$I(\theta) = E \left[ - \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta'} \right]$ ; matriz de información o de segundas derivadas para medir la convexidad.

Para el caso de un único parámetro, lo anterior se limitaría a:

$$W = (\hat{\theta} - \theta_0)^2 I(\hat{\theta}) \quad \text{Ec. 212}$$

Si se cumple la hipótesis nula (los verdaderos parámetros evaluados son cero), el valor de este estadístico se distribuye como una  $\chi_q^2$ , donde "q" es el número de parámetros accesorios que se propone en la hipótesis alternativa.

El cálculo del multiplicador de Lagrange se realiza a través de la expresión:

$$LM = \left( \frac{\partial L_T(\hat{\theta}_r)}{\partial \hat{\theta}_r} \right)' I(\hat{\theta}_r)^{-1} \left( \frac{\partial L_T(\hat{\theta}_r)}{\partial \hat{\theta}_r} \right) \Rightarrow \chi_p^2 \quad \text{Ec. 213}$$

donde  $L_T(\hat{\theta}_r)$  es el logaritmo de máxima verosimilitud evaluado para los parámetros estimados con el modelo restringido,  $I$  es la matriz de información (segunda derivada) y "p" es el número de restricciones que se imponen en la hipótesis nula (modelo restringido) frente a la alternativa (modelo ampliado).

La lógica del multiplicador de Lagrange parte del análisis de la pendiente de la función de verosimilitud para los valores de los parámetros propuestos. Dado que en el máximo de la función el valor de la pendiente debe ser nulo, se sigue inmediatamente que cuanto mayor sea la diferencia a cero en la pendiente calculada con otro valor paramétrico distinto al que hace máximo la función más evidente es que este parámetro no maximiza la función. Como se puede comprobar con la ilustración anterior, la pendiente en cada punto (para cada parámetro calculable), estará en función del grado de convexidad de la función, por lo que, para no obtener un contraste sesgado como ocurría con el ratio de verosimilitud, el

Multiplicador de Lagrange propone una ponderación por la medida de esta convexidad; es decir, la matriz de información o de segundas derivadas. Como en el caso anterior, los valores se toman al cuadrado para evitar la compensación de signos.

El Multiplicador de Lagrange se puede calcular como:

$$ML = S(\theta_0)' I(\theta_0)^{-1} S(\theta_0) \rightarrow \chi_q^2 \quad \text{Ec. 214}$$

donde,

$$S(\theta) = \frac{\partial \ln L(\theta)}{\partial \theta}; \text{ pendiente de la función en cada parámetro.}$$

$I(\theta) = E \left[ - \frac{\partial^2 \ln L(\theta)}{\partial \theta \partial \theta'} \right];$  matriz de información o de segundas derivadas para medir la convexidad.

Si se cumple la hipótesis nula (los verdaderos parámetros evaluados son cero), el valor de este estadístico se distribuye como una  $\chi_q^2$ , donde "q" es el número de parámetros accesorios que se propone en la hipótesis alternativa.

El contraste de Newey (1985) se puede considerar el caso específico del multiplicador de Lagrange para los modelos no lineales que contienen momentos condicionales. En muchas ocasiones la estimación de los parámetros de un modelo se fundamenta en que los resultados sean consistentes. Sin embargo, problemas de especificación deficiente (*misspecification*) darían lugar a estimadores de los parámetros no consistentes. En los modelos no lineales estimados por máxima verosimilitud la mala especificación es especialmente importante por producir importantes diferencias en los resultados en función de las asunciones hechas sobre las distribuciones.

El contraste de Newey incluye los de White (1984) y el de Hausman (1978) como casos específicos. Este contraste se fundamenta en la derivación de las distribuciones en el límite de diferentes contrastes de especificación para formas específicas de especificación deficiente.

Para el contraste de los momentos condicionales de un modelo en el que existen variables exógenas, Newey centra su propuesta en el poder informativo de la esperanza del vector objetivo (*score vector*) para una observación condicionado a las variables exógenas que, por definición, debe ser cero. De ello se deduce que las funciones de las variables exógenas deben estar incorrelacionadas con dicho vector de resultados. En definitiva, el contraste podría ser simplemente evaluar las funciones de covarianza entre las variables exógenas y el vector objetivo en el valor obtenido para el estimador que se pretenda.

White (1987) da lugar a los llamados contraste sobre la Matriz de Información Dinámica, que serán desarrollados por Perez Amaral (1993) y se basan en que sólo bajo el supuesto de especificación correcta las formas del hessiano y el producto exterior de la matriz de información son equivalentes.

### **Cambios de estructura en los modelos ARCH: contrastes de estabilidad de las varianzas**

Al igual que en los modelos econométricos estructurales, en los modelos tipo ARCH cobra pleno sentido realizar un contraste de la capacidad de dicho modelo para mantener la especificación dada a la evolución de la varianza durante el tiempo. Con este fin se pueden contrastar las siguientes hipótesis nula y alternativa:

$$H_0 = \theta_1 = \theta_2$$

$$H_a = \theta_1 \neq \theta_2$$

donde  $\theta_1$  y  $\theta_2$  son los parámetros estimados en una primera submuestra del total y en una segunda respectivamente. Como viene ocurriendo tradicionalmente en este tipo de contrastes de estructura, cabe distinguir dos casos: aquél en el que se conoce el punto en el que se produce el cambio de estructura y aquél en que dicho punto es desconocido. Shiang (1995)<sup>1</sup> desarrolla estadísticos para el contraste de ambos casos: punto de ruptura conocido y desconocido.

A) Conocido el punto de ruptura

La muestra total "T" quedaría dividida en dos partes de  $\delta T$  y  $T - \delta T$  observaciones cada una de ellas (siendo  $0 < \delta < 1$ ). La función de máximoverosimilitud se puede particionar del siguiente modo:

$$L(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T, h_t, \delta) = T^{-1} \sum_{t=1}^{\delta T} \log f_{1t} + T^{-1} \sum_{t=\delta T+1}^T \log f_{2t} \quad \text{Ec. 215}$$

donde

$$\log(f_{it}) = \frac{1}{2} \left( -\log(h_{it}) - \frac{\varepsilon_{it}^2}{h_{it}} \right), \quad h_{it} = \theta_i' Z_t \quad \text{Ec. 216}$$

y

$$\begin{aligned} \theta_i &= (\alpha_{i0}, \dots, \alpha_{iq}, \beta_{i1}, \dots, \beta_{ip}) \\ Z_t &= (1, \varepsilon_{t-1}^2, \dots, \varepsilon_{t-q}^2, h_{t-1}, \dots, h_{t-p}) \end{aligned} \quad \text{Ec. 217}$$

Empleando los valores de la función de verosimilitud para los parámetros estimados del modelo  $\theta_1$  y  $\theta_2$ , las condiciones de primer orden para la maximización de la función de verosimilitud son las siguientes:

$$\begin{aligned} \frac{\partial L}{\partial \theta_1} &= T^{-1} \sum_{t=1}^{\delta T} \frac{1}{2} h_{1t}^{-1} \left( \frac{\partial h_{1t}}{\partial \theta_1} \right) \left( \frac{\varepsilon_{1t}^2}{h_{1t}} - 1 \right) \\ \frac{\partial L}{\partial \theta_2} &= T^{-1} \sum_{t=\delta T+1}^T \frac{1}{2} h_{2t}^{-1} \left( \frac{\partial h_{2t}}{\partial \theta_2} \right) \left( \frac{\varepsilon_{2t}^2}{h_{2t}} - 1 \right) \end{aligned} \quad \text{Ec. 218}$$

Con el multiplicador de Lagrange antes comentado, se puede establecer un contraste de estabilidad de la varianza como:

$$LM(\delta) = T \left[ \frac{\partial L(\delta)}{\partial \theta'} \right] B_0^{-1} \left[ \frac{\partial L(\delta)}{\partial \theta'} \right]' \Rightarrow \chi^2_{p+q+1} \quad \text{Ec. 219}$$

donde la hipótesis nula es la de la estabilidad ( $\theta_1 = \theta_2 = \theta_0$ ) y el número de restricciones que en este caso se está suponiendo es que basta con una única estimación de los parámetros para toda la muestra (con lo cual no es necesario realizar la segunda estimación de  $p+q+1$  parámetros diferentes).

La matriz de información sería la siguiente:

$$B_0^{-1} = \frac{1}{4} \begin{bmatrix} \delta v_0 & 0 \\ 0 & (1 - \delta v_0) \end{bmatrix} \quad \text{Ec. 220}$$

donde  $v_0$  se define como:

$$v_0 = h_{0t} \left( \frac{\partial h_{0t}}{\partial \theta_1} \right) \left( \frac{\mathcal{E}_{t-1}^2}{h_{0t}} - 1 \right), \quad h_{0t} = \theta_0' Z_t \quad \text{Ec. 221}$$

Se acepta la hipótesis alternativa (existe un cambio de estructura) cuando el valor así calculado supera la chi-cuadrado con  $p+q+1$  grados de libertad.

#### B) Desconocido el punto de ruptura

La distribución del estadístico  $LM(\delta)$  es la función Bessell de orden "r". El multiplicador de Lagrange se escribiría, para este caso como:

$$LM1(\delta) = T \left[ \frac{\partial L(\delta)}{\partial \theta'} \right] (-A_0^{-1}(\delta)) \left[ \frac{\partial L(\delta)}{\partial \theta'} \right]' \Rightarrow B_r^2 \quad \text{Ec. 222}$$

donde se puede estimar de forma consistente:

$$\hat{A}_T(k) = \frac{1}{2} \begin{bmatrix} \frac{k}{T} \hat{G}_T & 0 \\ 0 & \left(1 - \frac{k}{T}\right) \hat{G}_T \end{bmatrix} \quad \text{Ec. 223}$$

$$\hat{G}_T = \frac{1}{T} \sum_{t=1}^T h_t^2(\hat{\theta}) \left[ \frac{\partial h_t(\hat{\theta})}{\partial \theta} \right] \left[ \frac{\partial h_t(\hat{\theta})}{\partial \theta} \right]'$$

Todo ello bajo el supuesto de que la especificación de la varianza es correcta, por lo que los residuos estimados se distribuyen como una normal u la varianza marginal es una martingala en diferencias ( $E_{t-1}(v_t) = 0$ ). En esta expresión,  $k$  es un número entero que se identifica con el punto de corte.

Para el caso en que la varianza no esté bien especificada, Shiang (1995) ofrece otras especificaciones más complejas en las que no entramos.





## **8. MODELIZACIÓN TIPO ARCH APLICADA AL IBEX-35**

## 8.1 Introducción y justificación del estudio

Los modelos que se han analizado durante la etapa teórica de esta tesis se han demostrado claramente útiles en la estimación de la varianza condicional de una serie de alta frecuencia. La justificación estadística del empleo de la modelización tipo ARCH está intrínseca en la determinación de un Proceso Generador de Datos que, de otro modo, se realizaría de forma incompleta, ya que se obtendría un "generador falsamente estacionario", o, dicho de otra forma, estacionario sólo de un modo marginal, ya que la varianza de la serie no se puede considerar constante en términos condicionales cuando la información de  $t-1$  es perfectamente conocida en  $t$ .

A partir de lo dicho, es necesario recurrir a una estimación de esta varianza condicional para la buena identificación de un proceso cuya estructura de autocorrelaciones en los valores absolutos o potencias (cuadradas o no) presenta valores significativos. Por supuesto, para la estimación paramétrica el conocimiento exacto de la verdadera varianza de la distribución que se intenta maximizar aumentará la eficiencia de los coeficientes calculados.

Centrándonos en el caso concreto que nos ocupa, la serie estadística que se empleará en la aplicación práctica es de orden financiero y, por ello, contiene una difícil apreciación del riesgo que contribuye de forma clara a su formación y cuyo conocimiento supone el mayor reto de los agentes económicos dedicados a la economía financiera desde que Bachellor en 1900 enunciara su principio de valoración de las opciones en su tesis doctoral.

La aplicación de los modelos tipo ARCH a la modelización financiera es algo común desde la creación de los mismos, pero en el momento actual, una buena estimación de los rendimientos de un índice como el Ibex-35 se hace especialmente interesante.

Desde la creación de la Unión Monetaria Europea con once países el 1 de enero de 1999, la política monetaria se unifica en manos del Banco Central Europeo. En el camino previo de convergencia para los países que integran el Euro, el descenso de la rentabilidad libre de riesgo ha sido el más vertiginoso conocido históricamente y la pérdida de atractivo de

la Deuda Pública, a tan bajo nivel de rentabilidad, se transmitió contundentemente a los mercados de renta variable como efecto equilibrador: ante la caída de la rentabilidad del mercado sin riesgo, la gran masa de inversionistas se traslada hacia el mercado de renta variable, con el consiguiente sobrepeso de la demanda de activos sobre su oferta y el notabilísimo "boom" bursátil que venimos viviendo en los últimos dos o tres años.

Las bolsas mundiales presentan cifras record de capitalización y la naturaleza de los agentes que en ella intervienen no puede ser más diversa. "El mercado se ha popularizado" con lo que ello tiene de cambio:

- Por un lado, la bolsa es más que nunca un juego en el que no sólo intervienen profesionales. Sus vaivenes (o su volatilidad) deben ser por tanto más frecuentes. La evolución de la bolsa es un fenómeno que interesa a multitud de observadores que ven como se modifica diariamente el valor de su ahorro y reaccionan ante estímulos diferentes a los que lo haría un agente financiero tradicional.
- Por otro lado y abundando al mismo tiempo en lo anterior, el conocimiento de las expectativas bursátiles y sus mecanismos generatrices es una cuestión de extrema relevancia. Los modelos de medición de la volatilidad como uno de los elementos más importantes en la medición de expectativas es un punto primordial.
- Finalmente, la sociedad de la información centra un importante número de actividades en la información precisa del movimiento de los activos financieros que hoy, más que nunca, gozan de los primeros lugares en las noticias, la radio, los teletextos, la televisión por cable, los paneles publicitarios electrónicos en la calle y, por supuesto, internet. No sólo la información es constante; la operatividad y acceso rápido a los mercados financieros está al alcance de cualquier ordenador personal o de cualquier teléfono. El número de operaciones se incrementa a ritmos inimaginables hace apenas unos años.

Tradicionalmente, la volatilidad en los mercados financieros se ha venido midiendo con el valor de la varianza del precio de los activos. Esta variable, la varianza, se convierte en explicativa de la evolución del propio precio en cuanto a información disponible en  $t$  de  $t-1$ . En cada momento del tiempo, la volatilidad conocida hasta el momento inmediatamente

anterior se convierte en un importante mecanismo para conocer la marcha del activo. Es un "hecho estilizado" suficientemente contrastado que los mercados financieros sufren de "contagio": a períodos muy volátiles les siguen períodos de igual característica y, una vez se "tranquilizan", la tendencia se mantiene en las observaciones ulteriores.

Por supuesto, medir la importancia de la volatilidad de forma aislada parece que va contra la propia naturaleza de cualquier variable económica, que casi nunca es fruto de una sola causa, sino de varias. En este contexto, el modelo de Glosten y otros (1993) parece indicado para determinar bien la volatilidad que luego será explicativa de la evolución del activo. Por otro lado, las aportaciones de Ding (1993) con su modelo A-PARCH parecen encajar bien con el comportamiento de los índices bursátiles.

En esta investigación, se propone el empleo de un modelo mixto fruto de las diversas aportaciones metodológicas que se han ido analizando previamente. Con el fin de contrastar las capacidad de los diferentes modelos e ir corrigiendo e incorporando nuevas formas de completar la incompleta especificación inicial, se irán aplicando las distintas modificaciones al modelo introducido por Engle en 1982. Con este proceso, se pretende lograr la especificación más adecuada para el Ibex-35, así como juzgar, en el terreno de la práctica, las posibilidades de los modelos analizados.

A partir de los contrastes teóricos y prácticos que se desarrollarán, se llegará a la conclusión que da nacimiento a una nueva forma de aproximar los modelos de volatilidad condicional, a la que denominaremos GJR-APARCH. También, durante las fases previas a lograr el modelo definitivo, se irán contrastando las diferencias en el empleo de una u otra forma funcional en la distribución, así como la "persistencia en volatilidad" y sus posibles correcciones a partir de modelos tipo IGARCH o GARCH de componentes.

## 8.2 Características básicas de la serie a modelizar

El IBEX-35 es el índice selectivo de la bolsa de Madrid. Su valor (precio) se calcula como media ponderada del precio de cada una de las acciones de las empresas que lo componen que, lógicamente, ha venido cambiando en los diez años de observaciones

diarias que se tienen en cuenta en esta aplicación debido a desapariciones de empresas, fusiones, cambio de nombre, etc.

La variación en las componentes del índice ha sido la que se puede ver en el cuadro adjunto para los últimos diez años.

EVOLUCIÓN DE LA COMPOSICIÓN DEL IBEX-35 HASTA JUNIO DE 1999						
COMPOSICIÓN ENERO 1991	PERÍODO	ENTRAN	SALEN	PERÍODO	ENTRAN	SALEN
ACE: Autopistas Conc. Española	ENE-JUN 91	AGR	AZU	JUL-SEP 91	MVC	ALB
ACX: Acerinox		CUB	PSG		CAN	CRI
ALB: Corporación Fin. Alba	OCT-DIC 91	HHU	PMD		PMD	ACX
ASL: Asland		ALB	HID	-	-	-
AZU: Azucarera	ENE-JUN 92	CRI	PMD	JUL-DIC 92	PRY	EXT
BBV: Banco Bilbao Vizcaya		SAR	CEN		ARA	HHU
BKT: Bankinter	ENE-JUN 93	EXT	HIS		ALB	AGS
BTO: Banesto		AGS	ECR	JUL-DIC 93	AUM	CEP
CEN: Banco Central	ENE-JUN 94	BCH	ALB		ARG	CEP
CEP: Cepsa		ACX	CRI	JUL-DIC 94	CTE	EBA
CRI: Cristalería Española	ENE-JUN 95	CEP	AGR		AGS	CTF
CTG: Catalana de Gas		HHU	ARA	JUL-DIC 95	AMP	GES
DRC: Dragados y Construcciones	ENE-JUN 96	AGS	URB		AZC	VDR
ECR: Ercros		ZOT	SAR	JUL-DIC 96	-	-
ELE: Endesa	ENE-JUN 97	EBA	ASL		AZC	ANA
FEC: Fecsa		GES	ZOT	JUL-DIC 97	TUB	GES
FOC: Focsa	ENE-JUN 98	ENC	HHU		VIS	MVC
HID: Hidrola		CTF	AGS	JUL-DIC 98	ACR	ALB
HIS: Bco. Hispano	ENE-JUN 99	GES	AZC		TPZ	FEC
IBE: Iberduero		UNI	ENC	JUL-DIC 99	FER	AMP
MAP: Corporación Mapfre	MZO-JUN 99	SOL	VIS		IDR	RAD
PMD: Petromed		UNI	VIS	MZO-JUN 99	SCH	RAD
POP: Banco Popular	MZO-JUN 99	SCH	BCH		SAN	
PSG: Prosegur		RAD	SAN			
REP: Repsol						
SAN: Banco Santander						
SEV: Sevillana de Electricidad						
TAB: Tabacalera						
TEF: Telefónica						
UNF: Unión Fenosa						
URA: Uralita						
URB: Urbis						
VAL: Vallehermoso						
VDR: Portland Valderrivas						
VIS: Viscofán						

ACR ACERALIA; ACS ACS.CONST; AGS AG.BARCELONA; ALT ALTADIS; AMP  
 AMPER; AMS AMADEUS; ANA ACCIONA; AZC;BBVA BBVA;BCH BANCO  
 CENTRAL HISPANO;CAN H.CANTABRICO;CTE CONTINENTE;CTF;CUB  
 ;EBA ;ENC ;EXT BANCO EXTERIOR;FCC FCC;FER G.FERROVIAL;GES  
 ;HHU ;HIS ;IBE IBERDROLA;IDR INDRA;NHH NH HOTELES;PRY  
 PRYCA;RAD ;REE RED ELEC.ESP.;SAR ;SCH BSCH;SGC  
 SOGECABLE;SOL SOL MELIA;TPI TPI;TPZ TELE PIZZA;TRR TERRA;TUB  
 ;UNI UNILEVER;ZEL ZELTIA;ZOT

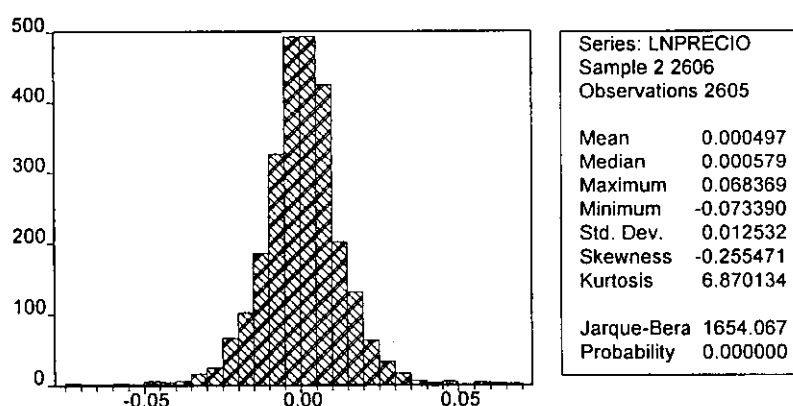
Fuente: Sociedad de Bolsas. Madrid

La serie empleada es el precio de cierre diario del IBEX-35 entre el 2 de enero de 1990 y el 5 de junio de 2000 para los días en los que la bolsa de Madrid tuvo actividad, con lo que será frecuente encontrar diferencias en el número de días transcurridos entre dos observaciones consecutivas (fin de semana, festivos, circunstancias especiales, ...). También es importante reseñar que el número de días hábiles de cada año es variable y oscila entre 248 y 251 según el período elegido.

Para aproximar el valor del rendimiento diario se emplea la diferencia de logaritmos del siguiente modo:

$$y_t = \text{LNPRECIO}_t = \text{LN}(\text{PRECIO}_t) - \text{LN}(\text{PRECIO}_{t-1}) \quad \text{Ec. 224}$$

Con las 2605 observaciones disponibles, se puede representar, a modo de resumen, los principales estadísticos descriptivos, que serían los siguientes:



De la representación del histograma de frecuencias se deduce sencillamente la presencia de una serie con un fuerte nivel de apuntamiento (curtosis) que supera con creces el valor de la normal: 6.87 frente al típico 3. Sin embargo, el valor de asimetría es muy próximo a cero (-0.25), lo que da lugar a una representación muy similar a la de la Normal, aunque mucho más apuntada.

Las características de esta serie están en línea con las que describieran Maldebrot y Fama (1965 y 1963) como comunes para las series de tipo financiero, que serían las siguientes:

- Carácter leptocúrtico de las observaciones
- Gruesas colas de simetría,
- Fuerte concentración en torno a la media

Uno de los puntos previos a la modelización tipo ARCH de las series es contrastar que nos encontramos ante un proceso estacionario en media. Para ello, seguimos el procedimiento apuntado por Dolado y otros (1990)<sup>ii</sup> y Perron (1989)<sup>iii</sup>.

En primer lugar se procede a realizar una regresión de la serie con una diferencia regular en función de un término constante, de ella con un retardo y de una variable tendencial.

En nuestro caso, la variable objetivo es LNPRECIO y su diferencia regular D(LNPRECIO), modelizada en función de su primer retardo (LNNPRECIO(-1)), una variable tendencial (@trend) y un término independiente. Los resultados de la regresión MCO son los siguientes:

Variable Dependiente: D(LNPRECIO)

Método: Mínimos Cuadrados Ordinarios

Muestra (ajustada): 2 2108

Observaciones incluidas: 2107 después de ajustar los extremos

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNPRECIO(-1)	-0.886058	0.021649	-40.92818	0.0000
C	0.000229	0.000547	0.417618	0.6763
@TREND(1)	3.34E-07	4.50E-07	0.742618	0.4578
Log Verosimilitud	6235.205	F-statistic		837.5581
Durbin-Watson stat	1.985280	Prob(F-statistic)		0.000000

Dada la no significatividad estadística de la variable de tendencia (la t-student experimental de @trend es 0.41, con lo que se acepta con un 67% de probabilidades que esta variable no es significativa para explicar la endógena) se estima un nuevo modelo sin incluir a ésta:



Variable Dependiente: D(LNPRECIO)

Método: Mínimos Cuadrados Ordinarios

Muestra (ajustada): 2 2108

Observaciones incluidas: 2107 después de ajustar los extremos

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNPRECIO(-1)	-0.885777	0.021643	-40.92581	0.0000
C	0.000580	0.000274	2.119169	0.0342
S.E. of regression	0.012555	Akaike info criterion		-5.916402
Sum squared resid	0.331822	Schwarz criterion		-5.911036
Log Verosimilitud	6234.929	F-statistic		1674.922
Durbin-Watson stat	1.985291	Prob(F-statistic)		0.000000

El valor de la t-student de la variable LNPRECIO(-1) continúa siendo altamente representativo de un valor no nulo del parámetro, por lo que, en principio, rechazaríamos la hipótesis nula de existencia de una raíz unitaria que se manifestaría en un valor cero del parámetro. Con este contraste intentamos comprobar si, existiendo una raíz unitaria, esta quedaría eliminada al practicar una diferencia, luego al regresar la diferencia de la endógena sobre ella misma sin diferenciar, el valor del parámetro debiera ser cero, caso de existir la raíz, ya que este parámetro estaría representado por  $\rho - 1 = 0$ ).

En la literatura está plenamente contrastada la sensibilidad del esquema propuesto ante la existencia de autocorrelación en los valores de la variable objetivo. A juicio del valor de Durbin Watson obtenido (1,98), no parece probable la existencia de autocorrelación; pero, aún así, y tal y como recomiendan diversos autores, se procede a incluir como explicativa la propia endógena desplazada.

Variable Dependiente: D(LNPRECIO)

Método: Mínimos Cuadrados Ordinarios

Muestra (ajustada): 2 2108

Observaciones incluidas: 2107 después de ajustar los extremos

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNPRECIO(-1)	-0.932726	0.028972	-32.19404	0.0000
D(LNPRECIO(-1))	0.053830	0.021782	2.471328	0.0135
C	0.000617	0.000274	2.250552	0.0245
S.E. of regression	0.012541	Akaike info criterion		-5.918207
Sum squared resid	0.330752	Schwarz criterion		-5.910155
Log Verosimilitud	6234.872	F-statistic		839.9809
Durbin-Watson stat	2.002480	Prob(F-statistic)		0.000000

Incluido este proceso autorregresivo en la ecuación, las conclusiones antedichas se mantienen, y, juzgar por los criterios informativos de Akaike o de Schwarz, el aporte a la explicación de la endógena de la inclusión de este retardo es mínimo, por lo que se opta por no incluir más variables explicativas de este tipo y concluir definitivamente que, en lo referente a la detección de raíces unitarias en la parte regular de la serie LNPRECIO, el contraste ampliado de Dickey Fuller es concluyente en rechazar la presencia de raíces unitarias con un 99% de probabilidad.

ADF Test Statistic	-32.19404	1% Critical Value*	-3.4365
		5% Critical Value	-2.8634
		10% Critical Value	-2.5678

Parece haber un consenso generalizado sobre la estructura tipo ARIMA de una serie financiera y, en particular, bursátil. Las numerosísimas aplicaciones que se han realizado sobre el tema parecen confirmar la existencia de un proceso de dilatada memoria que comprendería un número enorme de retardos en el tiempo. También es práctica generalizada abordar este proceso autorregresivo de orden "infinito" como su inversa matemática; es decir, como un proceso de medias móviles de orden uno. Inicialmente, y al igual que lo hicieran Engle (1986), Campbell (1992), Schwert (1989), etc., proponemos realizar la estructura el modelo tipo ARCH a partir del residuo generado por la aplicación

de un modelo MA(1) aplicado a nuestra serie de la diferencia de orden uno del logaritmo del precio diario del Ibex-35<sup>19</sup>.

El resultado de la regresión practicada se puede resumir del siguiente modo:

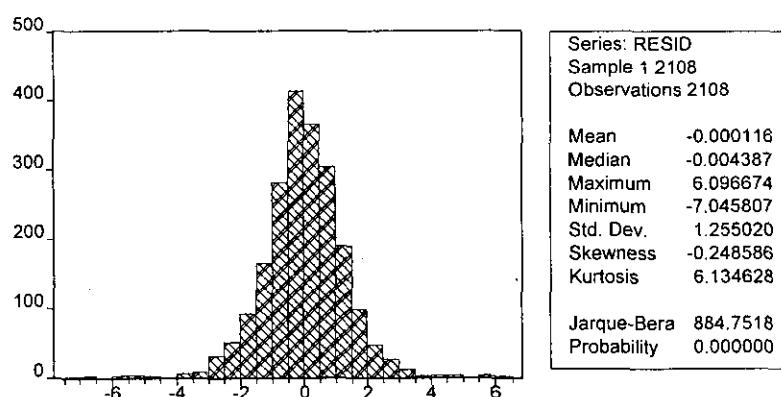
$$\text{LNPRECIO} = 0.06679327279 + 0.1246258717 \varepsilon_{t-1} + \varepsilon_t$$

(2,1724)                      (5,7616)

D-W: 2.006401                      Log Máx. Verosimilitud: -3464,457

El valor del estadístico L-jung-Box (descrito en el capítulo de contrastes de estructura y bondad de ajuste) nos ofrece un resultado de 55,56 para treinta seis retardos, lo que comparado con el valor de la  $\chi^2$  correspondiente nos permite afirmar que estamos ante un ruido blanco con un 97,9% de probabilidades.

Fruto de la modelización descrita nos encontramos con una serie del residuo que responde a las siguientes características:



Nuevamente, claramente leptocúrtica (el valor de curtosis sobrepasa el 6) y ligeramente asimétrica (cargada hacia los valores negativos con un coeficiente del -0.24). El valor del estadístico Jarque-Bera no deja lugar a dudas sobre la no normalidad de la serie planteada.

<sup>19</sup> Como es habitual en todos los documentos empleados para la elaboración de esta tesis, aquí también se utiliza el valor de la diferencia del logaritmo del precio multiplicado por 100. La reducida proporción de la resta de un valor que ya está en logaritmos origina valores tan pequeños que, salvo con esta pre-multiplicación, son intratables matemáticamente con un ordenador.

### 8.3 Estructura muestral de las funciones de autocorrelación de la serie para diferentes transformadas de la misma: presencia de estructura ARCH en los residuos

Como es obvio, el primer paso para la posterior modelización que se propone en esta tesis es comprobar si realmente existe o no una estructura tipo ARCH en la serie temporal del Ibex-35 durante los últimos nueve años. En cada una de las especificaciones tipo ARCH que se han ido repasando y, en particular, en el capítulo correspondiente al contraste del modelo se abarcaron las diferentes formas para el chequeo de la existencia de una estructura tipo ARCH.

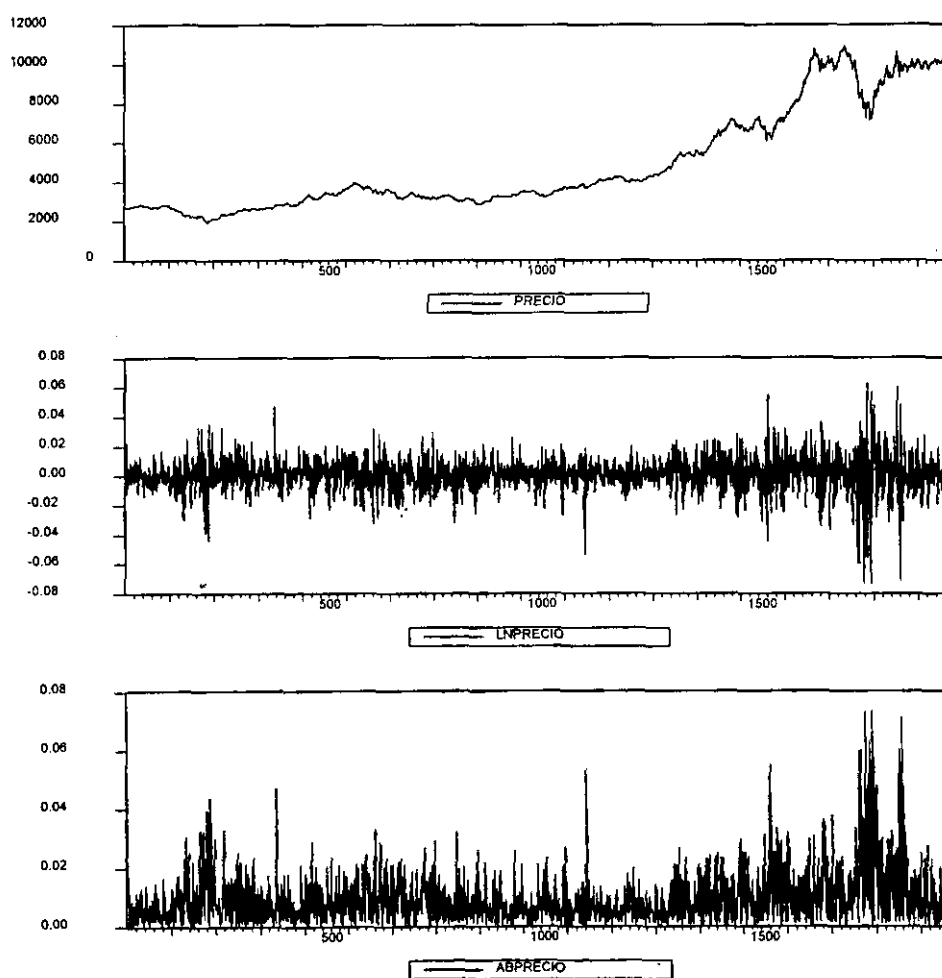
En particular, se han venido poniendo de manifiesto, desde los primeros artículos referidos a este tipo de modelos, que, si bien puede presentarse un proceso como estacionario en términos marginales, en algunos casos un examen detallado de momentos superiores de la serie que se esté tratando pueden manifestar grados de correlación significativos. De un modo concreto, Engle y Bollerslev proponen la contrastación de los valores muestrales al cuadrado, por ser estos los que se corresponderían con el valor de la varianza; pero también en valores logarítmicos y valores absolutos.

Como ya ha sido referido con suficiente detalle, Ding (1993) se cuestiona también el valor del momento asimétrico que manifiesta autocorrelación llegando a la conclusión de que, en grado máximo, este no es el del cuadrado, sino el de una potencia que oscila entre 1 y 1,5 en las diferentes evidencias empíricas sobre las que trabaja. En este apartado, seguiremos el esquema de contrastación llevado a cabo por este autor en su caso para la serie del S&P 500 y, en el nuestro, para el Ibex-35.

El gráfico siguiente muestra la evolución de las series del precio diario, del rendimiento diario calculado tal y como se expresaba en la Ec. 224 y del valor absoluto de este último. Ante estas representaciones, se puede concluir los mismos hechos que comenta Ding (1993) para el caso del índice S&P-500:

- El precio del índice (PRECIO) manifiesta una clara tendencia creciente si se observa en todo el recorrido presentado.

- Observando la representación del rendimiento (LNPRECIO), dicha tendencia no existe o el valor se distribuye de manera estable en torno a la media cero (0.000711)
- En el último gráfico, el valor absoluto del rendimiento (ABPRECIO), pone de manifiesto las ya hartamente citadas conclusiones de Fama (1965) y Malmelbrot (1963) sobre el carácter de contagio en la intensidad de los movimientos: a períodos altamente volátiles le suceden períodos de iguales características y viceversa.



Tal y como se ha venido poniendo de manifiesto teóricamente, los valores absolutos de los rendimientos manifiestan una mayor autocorrelación que los valores normales. Este hecho también se repite para retardos superiores a tres en el caso de los cuadrados de las series.

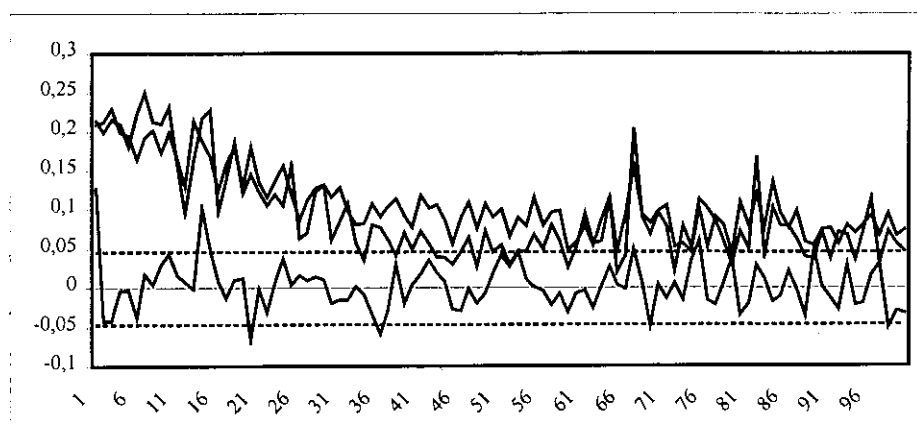
Los resultados que figuran en la tabla siguiente parecen concordar con el estudiado proceso de autocorrelación de orden uno que diversos autores estiman en los rendimientos financieros (Fama (1976), Taylor (1986), Hamao (1990), entre otros muchos). En palabras de Ding (1993), el valor del primer retardo (0.127) representa este valor del autorregresivo de orden uno. El valor negativo de la autocorrelación de segundo orden podría representar el comportamiento de recuperación de beneficios típico de las variables financieras llamado, en diversos estudios, reversión a la media (*mean reversion*). El primero de estos puntos es significativamente distinto de cero a juzgar por la bandas de confianza para el 95% de probabilidad calculadas como  $\pm 1.96/\sqrt{T}$ <sup>20</sup>. En el segundo, el hecho no es tan claro.

Autocorrelaciones del rendimiento, su valor absoluto y su cuadrado

	1	2	3	4	5	10	20	40	70	100
LNPRECIO	0,127	-0,043	-0,043	-0,003	-0,002	0,046	-0,067	0,003	0,003	-0,034
ABPRECIO	0,209	0,212	0,229	0,199	0,194	0,198	0,180	0,077	0,099	0,074
LNPRECIO2	0,214	0,199	0,215	0,209	-0,181	-0,231	0,145	0,049	0,096	-0,047

<sup>20</sup> Según Barlett (1946), el proceso de autocorrelación  $\rho_1$  se distribuye aproximadamente como una normal  $N(0,1; T)$ . En el caso expuesto, las bandas serían  $\pm 0.0465$  y se representan gráficamente en líneas discontinuas.

## Representación gráfica del cuadro anterior



Para el caso de las autocorrelaciones en los valores absolutos y cuadrados de las serie de rendimientos, se cumplen las dos circunstancias señaladas por Ding (1993): no solo son claramente significativos distintos de cero sino, además, casi siempre en valores positivos. El hecho que no concuerda con la presunción de Ding es el de que la serie de rendimientos absolutos es siempre mayor que el de la serie de rendimientos al cuadrado. En nuestro caso, existen muchos momentos en los que ocurre lo contrario, aunque hay que señalar que, en general, o bien son coincidentes o bien está ligeramente por encima la serie de valores absolutos.

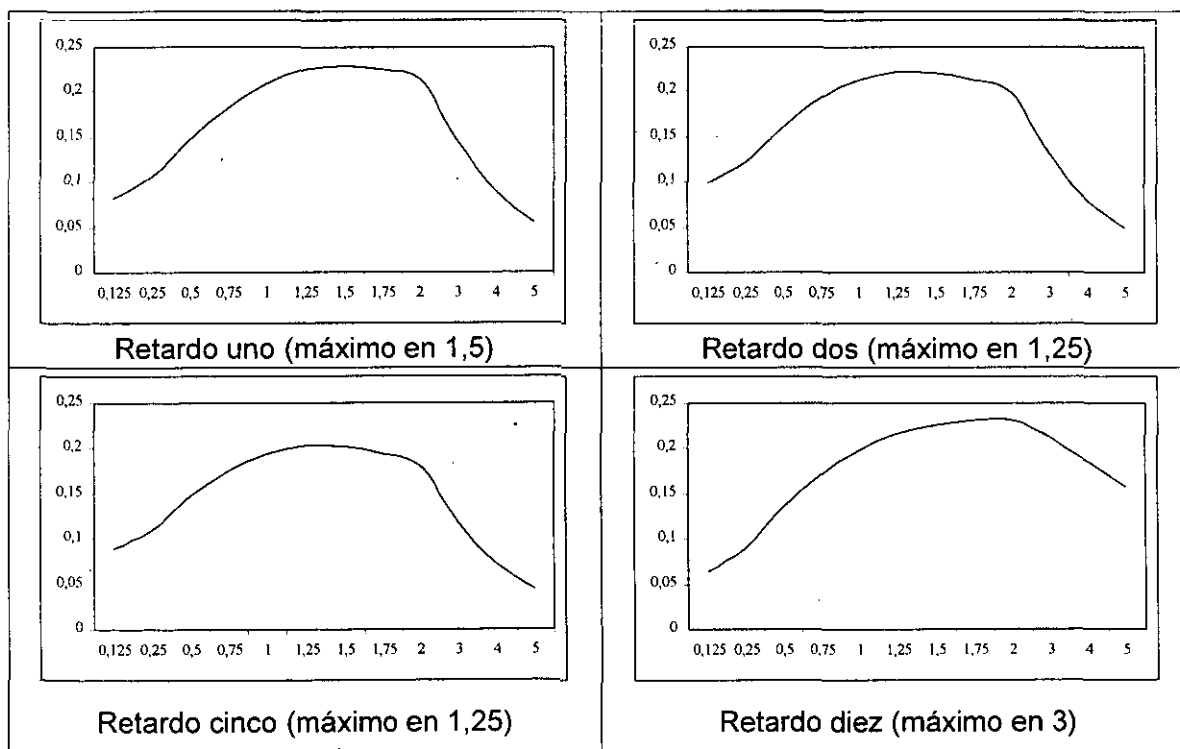
La situación contrastada hasta el momento parece poner de manifiesto la "memoria larga" de la serie considerada, por lo menos hasta el retardo 100. El hecho se produce también para los rendimientos absolutos elevados a diferentes potencias. En la tabla adjunta se muestran para potencias entre 0.125 y 3, y las conclusiones observadas hasta el momento se mantienen.

Autocorrelaciones del rendimiento absoluto elevado a "d"

D	1	2	3	4	5	10	20	40	70	100
0,125	0,082	0,099	0,134	0,092	0,089	0,064	0,088	0,033	0,026	0,030
0,25	0,105	0,121	0,154	0,111	0,110	0,089	0,108	0,045	0,039	0,040
0,5	0,147	0,161	0,188	0,146	0,147	0,135	0,142	0,064	0,064	0,057
0,75	0,182	0,192	0,214	0,176	0,175	0,171	0,166	0,074	0,085	0,069
1	0,209	0,212	0,229	0,199	0,194	0,198	0,180	0,077	0,099	0,074
1,25	0,224	0,222	0,235	0,213	0,203	0,216	0,183	0,075	0,106	0,073
1,5	0,229	0,221	0,233	0,219	0,202	0,226	0,177	0,068	0,107	0,066
1,75	0,225	0,213	0,226	0,217	0,194	0,231	0,163	0,059	0,103	0,057
2	0,214	0,199	0,215	0,209	0,181	0,231	0,145	0,049	0,096	0,047
3	0,145	0,131	0,168	0,153	0,117	0,213	0,072	0,016	0,065	0,014

Además de lo dicho hasta ahora, la autocorrelación presenta los valores más elevados cuando la potencia es uno o valores muy cercanos a uno.

Autocorrelación del rendimiento absoluto elevado a "d"



En eje de ordenadas están representados los valores de autocorrelación y en el de abscisas la potencia que se ha aplicado (el valor de "d").



En estos gráficos se ilustra la conclusión anterior: la autocorrelación de las potencias del valor absoluto del rendimiento muestran el valor mayor cuando dichas potencias son uno o valores muy próximos a uno. La función es, en ambas colas desde uno, casi monótonicamente decreciente. Esto es así para cualquier retardo analizado, aunque en el caso del retardo diez, la función no está tan claramente centrada en uno y los resultados no son tan espectaculares como los logrados por Ding para el índice S&P-500.

En el caso del IBEX-35, las autocorrelaciones del valor absoluto del proceso muestran algunos valores negativos en períodos mucho más cercanos a los referidos por dicho autor. En su caso, dichas observaciones negativas no se producían más que de forma casual cada diez años (como término medio). En el caso del índice español, se producen algunos puntos con correlación negativa cada año, aunque dichos puntos se pueden considerar anormales, dada su escasísima existencia (en las primeras 100 autocorrelaciones, sólo el valor de autocorrelación con el punto t-74 presenta un valor negativo).

Según este mismo autor, para contrastar una caída lenta al principio y cada vez más acelerada en momentos subsiguientes en la autocorrelación de las series financieras (hecho palpable en los cálculos muestrales realizados), se puede modelizar dicha autocorrelación según la siguiente función<sup>21</sup>:

$$\rho_t = \frac{\alpha \rho_{t-1}^{\beta_1} \beta_2^t}{t^{\beta_3}}$$

en la que tendencia autorregresiva de la correlación se reduce a partir de la expresión exponencial que aparece en el denominador a medida que transcurre el tiempo. Operativamente, se transforma en una función lineal tomando logaritmos:

$$\ln(\rho_t) = \ln(\alpha) + \beta_1 \ln(\rho_{t-1}) + t \ln(\beta_2) - \beta_3 \ln(t)$$

<sup>21</sup> Ding (1993) rechaza utilizar directamente los métodos empleados por Granger y Joyeux (1980) de integración fraccionada o el típico de un proceso ARMA porque, en ambas situaciones, el resultado está en contra de la realidad empírica expuesta: lenta caída al comienzo y, progresivamente, cada vez más rápida. El autor propone una función que es mezcla de las dos anteriores comentadas.

Realizando esta estimación con MCO, los resultados para el caso que nos ocupa son los siguientes<sup>22</sup>:

$$\text{LOG}(\rho_t) = -1.263 + 0.2310 \cdot \text{LOG}(\rho_{t-1}) - 0.009 \cdot t - 0.006 \cdot \text{LOG}(t)$$

(3.02)                      (4.79)                      (0.07)

$$R^2=0,70 \quad \text{DW}= 1,65$$

En este caso, los resultados obtenidos son menos sorprendentes que en de Ding, en el que la capacidad de estimación del modelo con esta función resultaba mucho más exacta.

Como ya es patente, el contraste de las funciones de autocorrelación para diversas transformadas de la serie de la diferencia del logaritmo del precio es claramente significativo de la presencia de una estructura tipo ARCH, ya que lo es su correlación para los valores al cuadrado, en valor absoluto y para diferentes potencias asimétricas. Dado que la modelización que se propone es sobre el residuo del modelo ya estimado mediante un proceso de medias móviles de orden uno, tiene sentido aquí probar que la estructura de dicho residuo también es de tipo ARCH.

Autocorrelaciones del residuo absoluto elevado a "d"

d	1	2	3	4	5	10	20	40	70	100
0,125	0,06	0,11	0,15	0,09	0,12	0,09	0,08	0,03	0,02	0,04
0,25	0,08	0,13	0,16	0,11	0,14	0,11	0,10	0,04	0,04	0,05
0,5	0,12	0,17	0,19	0,14	0,17	0,15	0,14	0,06	0,06	0,07
0,75	0,16	0,20	0,21	0,17	0,19	0,18	0,16	0,06	0,08	0,08
1	0,19	0,22	0,22	0,19	0,21	0,20	0,18	0,07	0,09	0,08
1,25	0,21	0,23	0,23	0,21	0,21	0,22	0,18	0,07	0,10	0,08
1,5	0,22	0,23	0,23	0,22	0,21	0,23	0,18	0,06	0,10	0,08
1,75	0,23	0,23	0,23	0,22	0,21	0,24	0,17	0,05	0,09	0,07
2	0,22	0,21	0,22	0,21	0,19	0,24	0,15	0,04	0,09	0,06
3	0,17	0,15	0,18	0,16	0,14	0,22	0,08	0,02	0,05	0,03
4	0,12	0,10	0,14	0,11	0,09	0,19	0,04	0,00	0,03	0,01
5	0,08	0,06	0,11	0,07	0,07	0,16	0,02	0,00	0,02	0,00

<sup>22</sup> La muestra empleada en esta regresión se refiere a los valores de la función de autocorrelación

Sobre los datos resumidos en la tabla anterior se pueden hacer tres observaciones:

- Existe correlación significativa para los valores absolutos de la serie y sus diversas potencias.
- La correlación promedio máxima se da en el valor absoluto elevado en el intervalo 1,5-1,75.
- Los valores de la autocorrelación para las distintas potencias del valor absoluto presentadas son siempre positivos hasta retardos muy lejanos y siempre que el valor de la potencia sea menor que dos. A partir de esta potencia, comienzan a aparecer valores de correlación negativos sobre el retardo 90 y siguientes.
- Los valores presentados son significativamente distintos ---de cero para retardos muy avanzados en el tiempo, haciendo esto incidencia sobre el tan comentado fenómeno de la "persistencia temporal" que se ha venido comentado sobre distintos momentos muestrales de las series financieras.

Realizando una regresión del residuo elevado al cuadrado sobre él mismo con un retardo, los resultados son los siguientes

$$e_t^2 = 0,000122 + 0,2210 e_{t-1}^2$$

(14,77)                      (10,4)

Obviamente la significatividad de la correlación (aquí medida por el valor de la t-student calculada con 10,4 y una probabilidad de parámetro nulo del 0%) nos invitan nuevamente a desarrollar una modelización tipo ARCH en el caso que nos ocupa.

Aplicando el contraste del multiplicador de Lagrange que sugiriera Engle para la determinación de una estructura ARCH, partimos de una muestra empleada de 2107 observaciones y la  $R^2$  obtenida en la regresión anterior es de 0.048881, por lo que su producto es 103,0411. Bajo la hipótesis nula validez del modelo restringido (sin incluir

---

muestral hasta el retardo 200.

retardo alguno), este valor debiera corresponder a una chi cuadrado con dos grados de libertad, hecho que el estadístico calculado rechaza con una probabilidad cercana al 100%.

## 8.4 Aplicación de diversas modelizaciones tipo ARCH a la serie diaria de rendimientos del ibex-35

Es frecuente en la literatura de los últimos años sobre los modelos ARCH recoger un elenco de posibles modelizaciones tipo ARCH para elegir, de entre ellas, aquella que es más robusta en virtud de su contrastación estadística. Tanto Pagan y Schwert (1990), como Engle y Ng (1993) o Glosten y otros (1993) - por recoger ejemplos relativamente recientes y que han marcado verdaderos hitos en la modelización de la varianza condicional - realizan un esquema de aproximación hacia el verdadero modelo a partir de la propuesta de un conjunto de submodelos cada vez más completo.

Replicaremos en este estudio dicho esquema a partir de las siguientes propuestas metodológicas:

### Modelos propuestos en Glosten y otros (1993)

Modelos	Especificación	Comentario
<b>Modelo I GARCH-M</b>	$y_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $h_{t-1} = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-2}$	GARCH en media con valor desplazado de la varianza y MA(1)
<b>Modelo II GARCH-M Ampliado</b>	$y_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $h_{t-1} = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2 i_{t-1} + \beta_1 h_{t-2}$	Modelo I + efecto apalancamiento (impacto de innovaciones pasadas negativas o positivas)
<b>Modelo III Shocks Asimétricos</b>	$y_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $h_{t-1} = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 i_t + \beta_1 h_{t-2} + \beta_2 r_f$	Modelo II + variable de activo sin riesgo
<b>Modelo IV Factores Estacionales</b>	$y_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $h_{t-1} = \alpha_0 + \alpha_1 \eta_{t-1}^2 + \alpha_2 \eta_{t-1}^2 I_{t-1} + \beta_1 h_{t-2}$ $\eta_t = (1 + \lambda_1 EST1_t + \lambda_2 EST2_t) \varepsilon_t$	Modelo III + Incorporación de los efectos estacionales típicos de los mercados bursátiles
<b>Modelo V Estructura E-GARCH</b>	$y_t = a_0 + a_1 h_{t-1} + \varepsilon_t$ $\log(h_{t-1}) = \alpha_0 + \alpha_1 \eta_{t-1}^2 / \sqrt{h_{t-2}} + \alpha_2 \eta_{t-1}^2 I_{t-1} / \sqrt{h_{t-2}} + \beta_1 \log(h_{t-2})$	Forma logarítmica (forzar valores positivos) y término EGARCH (romper la proporcionalidad en el impacto de innovaciones anormalmente altas)

Al esquema que proponían Glosten y otros (1993), incorporamos algunas variantes más recientes en la modelización tipo ARCH, así como algunas nuevas aproximaciones que cuya potencia se prueba en esta tesis con su utilización para la determinación de la volatilidad del Ibex-35. Dichas aportaciones serán:

- Un modelo IV alternativo (GJR-ARCH + ARCH de componentes), que incluirá un tratamiento especial de la "persistencia en volatilidad" modelizada a partir de un esquema de ARCH de componentes, discriminando una modelización a corto plazo y otra a medio y largo en la especificación de la volatilidad.
- Una especificación que recoja algunas características concretas de la estructura de autocorrelación del residuo, para determinar con un mayor grado de garantías cuál debe ser la potencia óptima del residuo para la estimación de la volatilidad. Con este objetivo, se plantea la estimación tipo APARCH (modelo VI) y una variante de esta que hemos venido a llamar GJR - APARCH (modelo VII).

**Modelo I: GARCH-M (1,1)<sup>23</sup> - la varianza es explicativa del modelo en medias con un retardo -**

El primer modelo que se presenta se corresponde a un GARCH (1,1) en media, dado que se ha especificado la serie del índice Ibex-35 con un proceso de medias móviles previamente. La especificación sería la siguiente:

$$\begin{aligned}\varepsilon_t &= y_t - \mu - \theta\varepsilon_{t-1} - \gamma h_{t-1} \\ h_t &= w + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1}\end{aligned}$$

donde  $y_t$  es la diferencia del logaritmo del precio diario del Ibex-35 tal y como ya se definió en diversas ocasiones. La modelización se realiza multiplicando este valor por 100 y este hecho se mantendrá para los modelos que se explicitan a continuación, manteniendo así la comparabilidad entre ellos. La única razón para realizar este producto se debe a necesidades de convergencia, ya que las ínfimas cifras que presenta la diferencia de un logaritmo son difícilmente tratables con el ordenador a no ser que se premultipliquen tal y como se ha dicho.

La elección de un modelo GARCH(1,1) y no otro de orden superior se debe a la contrastación de los valores de los parámetros calculados, que resultan no significativos para retardos superiores en la estimación de la varianza condicional autorregresiva. En principio, y según la estructura de la función de autocorrelación que se ha contrastado en el apartado de la estructura de la serie, la serie para el cuadrado del residuo presenta el correlograma típico de un ARMA(1,1), y sobre este esquema es sobre el que se inicia la especificación, comprobando además que retardos superiores tanto de la varianza condicional como del residuo precedente al cuadrado no resultan significativos en ningún caso.

Es importante poner aquí de manifiesto que el modelo planteado no es estrictamente un modelo en media del tipo descrito por Engle y otros (1986), ya que el valor de la varianza condicional como explicativa del proceso original se presenta aquí con un retardo, y no de

<sup>23</sup> Tal y como proponen los autores del modelo GJR-ARCH (1989), la varianza condicional funciona como explicativa del modelo en medias con un retardo; es decir  $h_{t-1}$  y no  $h_t$  es la variable que se utiliza en estos modelos.

forma contemporánea tal y como lo hicieran los referidos autores. Este hecho, común en las especificaciones recogidas sobre otros índices, parece corresponderse mejor con la realidad de la información disponible sobre la varianza condicional por el agente económico en el momento "t": parece que hay un "retardo" en la conexión de la información sobre la volatilidad por la no inmediatez de los cálculos en la propia jornada bursátil.

A continuación se muestran los resultados de la estimación máximo-verosímil que se ha realizado.

	Parámetros		Desviación Típica	Z- estadística	Prob.
	Nombre	Valor			
<b>Modelo de la Varianza condicional</b>	OMEGA	0.043369	0.008851	4.900021	0.0000
	ALPHA	0.093936	0.013254	7.087365	0.0000
	BETA	0.878140	0.016850	52.11426	0.0000
<b>Modelo de la Media</b>	MU	0.073629	0.045088	1.632991	0.0512
	THETA	0.118658	0.023508	5.047465	0.0000
	GAMMA	0.015331	0.035984	0.426053	0.3350
<b>Log Verosimilitud</b>		-3252.050	<b>Akaike info criterion</b>		3.09406
<b>Media Log Verosimilitud</b>		-1.544183	<b>Schwarz criterion</b>		3.11016
<b>Nº de explicativas</b>		6	<b>Hannan-Quinn criter.</b>		3.09996

Hay que señalar sobre este modelo algunas de las circunstancias comunes a otras variables de tipo financiero que han sido estudiadas en numerosas ocasiones:

- Tal y como ya se sabía por los contrastes de autocorrelación previamente mostrados, la estructura ARCH de la serie queda plenamente contrastada con la significatividad del proceso GARCH (1,1) que aquí se expone. El contraste de Wald confirma la no nulidad de los parámetros, tanto de forma conjunta como individual, con una probabilidad de prácticamente el 100%.
- La "positividad" de las estimaciones de la varianza condicional queda plenamente asegurada merced a que todos los parámetros estimados de la especificación de dicha varianza condicional han resultado positivos, con lo que queda asegurado



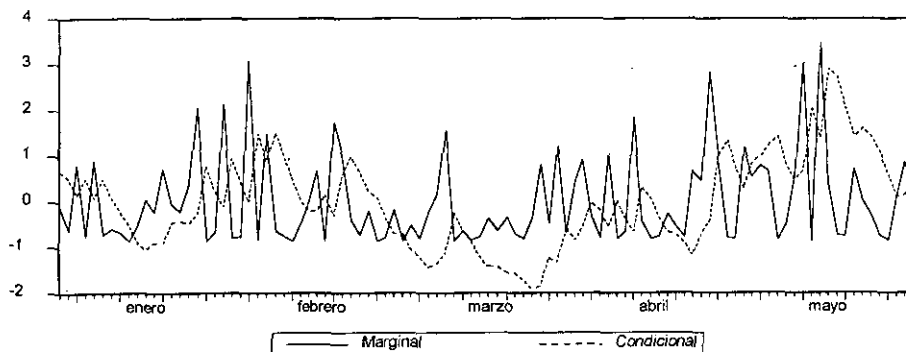
que su producto por números también positivos (el cuadrado del residuo y la varianza previa) darán siempre un resultado positivo en este modelo.

- La persistencia en la volatilidad es un hecho garantizado dado que la suma de los parámetros  $\alpha + \beta$  es prácticamente igual a uno (0,96). Este hecho, común a otras estimaciones realizadas sobre otros índices bursátiles o variable de precios, de tipos de interés o tipos de cambio, puede llegar a plantear graves problemas sobre la estacionariedad del proceso tal y como se ponía de manifiesto en el capítulo dedicado a los modelos IGARCH. En cualquier modo, también es cierto que una especificación deficiente del proceso ARCH puede dar lugar a este tipo de parámetros en suma cercanos a uno, hecho que se corregiría realizando una modelización adecuada<sup>24</sup>.
- En este modelo, el valor de la varianza condicional desplazada en la especificación de la media resulta no significativo, hecho que no concuerda con el principio básico de los modelos CAPM y de la lógica de los mercados financieros de que la volatilidad es un factor explicativo de la prima de riesgo. Esta circunstancia es muy relevante, por cuanto está en la base de nuestro punto de partida teórico el que debe haber una significatividad clara.

El gráfico de la serie de la varianza condicional estimada y de la varianza marginal (calculada directamente como cuadrado de los residuos esta última) muestra un cierto comportamiento tipo réplica de la marginal por la condicional con un retardo. Es decir; parece que la variable que se muestra significativa en la estimación del rendimiento del índice bursátil es un conocimiento aproximado de la varianza calculable en el período anterior más un cierto componente añadido por los errores que en este conocimiento se produjeron. Aún así, es obvio que el seguimiento de la varianza marginal por la condicional es deficiente y se hallan, en el gráfico, multitud de puntos en los que el signo de los puntos críticos es exactamente el contrario.

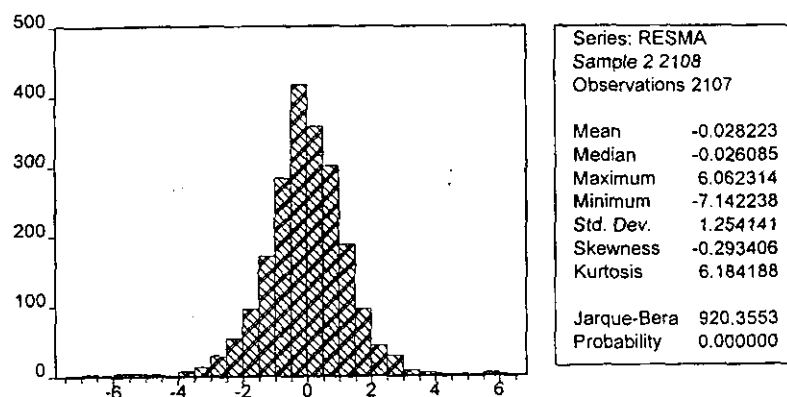
<sup>24</sup> Glosten y otros (1993) proponen también como medida aproximada de la "persistencia en volatilidad" el parámetro de un modelo AR(1) de la volatilidad estimada y su significatividad estadística medida por su correspondiente t-estadística. En el modelo 1 que aquí se presenta, el valor de dicho parámetro sería de 0.87, significativo con más del 99% de probabilidad; luego se podría concluir que la persistencia es muy grande.

Gráfico de la varianza marginal y condicional estimada en el año 2000



En el gráfico de los últimos períodos de las series de la varianza marginal y condicional (se muestra de enero a junio del año 2000), la aproximación de una serie por la otra es buena con el retardo comentado y salvo para el mes de marzo, donde las fuertes recogidas de beneficios marcaron cambios importantes en el día a día de las bolsas españolas, que se distenderían en los meses siguientes.

Los resultados del residuo estimado muestran con claridad una distribución no normal de dicha variable. Aún cuando la forma del gráfico es ligeramente parecida, nuevamente su nivel de apuntamiento es muy superior al de la normal (curtosis de más de 6, frente al típico 3 de la normal) y es algo asimétrica balanceada hacia la izquierda; es decir, hacia los valores negativos (coeficiente de asimetría de -0.29). Como resultado, el contraste de Jarque Bera rechaza sin lugar a dudas la posibilidad de una distribución normal de esta serie; lo que, en principio, era uno de los presupuestos sobre los que se asentaba el modelo ARCH y ha sido la función de densidad de una normal la que se utilizó para obtener los parámetros que anteriormente se comentaban.



Es obvio y ya ha quedado patente en multitud de ocasiones que la forma de la función de distribución de la serie dista bastante de ser una normal. Sin embargo, también se ha reflejado en la literatura y en diversos capítulos de esta tesis que multitud de experimentos de Montecarlo prueban la escasa ganancia que se deriva del empleo de la forma de una t-student (claramente más similar a esta que nos presenta el residuo estimado) que, sin embargo, introduce diversos problemas para llegar a la convergencia dado que incorpora un parámetro más a estimar ya que los grados de libertad son desconocidos<sup>25</sup>. A pesar de lo dicho y de cara a la comparación con el modelo propuesto y los parámetros estimados con la función normal, se ha generado un programa que emplee la t-student para ver si hay diferencias notables. Los resultados del modelo son en este caso los siguientes:

Parámetros			z-		Prob.
	Nombre	Valor	Desv. Típica	estadística	
<b>Modelo de la Varianza condicional</b>	OMEGA	0.031974	0.009004	3.551226	0.0002
	ALPHA	0.093697	0.015473	6.055642	0.0000
	BETA	0.887737	0.017283	51.36480	0.0000
<b>Modelo de la media</b>	MU	0.056190	0.042887	1.310199	0.0951
	THETA	0.115151	0.023039	4.998050	0.0000
	GAMMA	0.031668	0.032242	0.982216	0.1630
<b>Grados de libertad t-student</b>		10.49711	1.576556	6.658259	0.0000
<b>Log Verosimilitud</b>		-3227.613	<b>Akaike info criterion</b>		3.07180
<b>Media LogVerosimilitud</b>		-1.532580	<b>Schwarz criterion</b>		3.09059
<b>Nº de explicativas</b>		7	<b>Hannan-Quinn criter.</b>		3.07868

Hay que destacar que no existen grandes ventajas para emplear la estimación con la t-student en este caso, dado que, a juicio de diversos contrastes, no hay diferencias significativas ni en la reducción del error ni en los valores resultantes para los estimadores, que apenas difieren considerando la función de distribución normal o la t-student en la maximización. Recordemos que a similares conclusiones llegaban Engle y González-Rivera (1991) para índices diferentes al español y tras determinadas simulaciones de Montecarlo.

Volviendo entonces al modelo anterior, y de cara a comprobar la existencia de posibles efectos asimétricos en el modelo, aplicaremos los contrastes de sesgo de signo propuestos en Engle y Ng (1993). En primer lugar definimos una variable ficticia del siguiente modo:

$$S_{t-1}^- = \begin{cases} 1 & \text{si } \varepsilon_{t-1} < 0 \\ 0 & \text{resto casos} \end{cases}$$

Se plantean entonces los tres contrastes de Engle y Ng (1993):

<b>Contraste</b>	<b>Objetivo</b>	<b>Variable introducida</b>	<b>Resultado Obtenido<sup>26</sup></b>
Sesgo en el signo	existencia o no de asimetría	$S_{t-1}^-$	-1.43 (-2.85)
Sesgo negativo	efectos de "apalancamiento" ( <i>leverage</i> ) en valores negativos	$S_{t-1}^- \varepsilon_{t-1}$	0.054 (1.92)
Sesgo positivo	efectos de "apalancamiento" en valores positivos	$S_{t-1}^+ = (1 - S_{t-1}^-) \varepsilon_{t-1}$	-0.043 (1.27)

Los modelos contruidos a partir de la nueva ficticia<sup>27</sup> evidencian de forma sobradamente significativa la importancia del signo de la innovación (residuo) previo para la realización de la actual varianza. En definitiva, ponen de manifiesto que existe un efecto de

<sup>25</sup> ver Nelson (1991)

<sup>26</sup> Parámetro y z-estadística.

<sup>27</sup> Los modelos se especifican incluyendo estas variables como explicativas del residuo estandarizado tal y como proponen Engle y Ng (1993). Ver capítulo de contrastación de estructura para más detalles.

generación de mayor volatilidad cuando el período precedente "sorprendió" con un error negativo que cuando lo hizo de otra forma (a la luz del segundo y tercer contraste, de los cuales es el segundo significativo casi al 95%). En cualquier caso, la situación exige una modelización que incluya esta variable o algún otro modo de incluir el efecto "apalancamiento".

## Modelo II: Modelo I + efecto apalancamiento - variable "i" -

Dadas las conclusiones que se pueden realizar sobre los contrastes de sesgo de signo, parece oportuno introducir en la especificación de la varianza un elemento que discrimine la volatilidad calculada en función de si esta proviene de un período en el que la estimación que se realizó estuvo infravalorada o sobrevalorada.

Al margen de la evidencia estadística de los resultados mostrados, parece lógica la consideración de que sobrevaloraciones en la rentabilidad del Ibex-35 son rápidamente fruto de reacción para evitar que se disparen las pérdidas. Sin embargo, el aumento de la rentabilidad por encima de lo esperado (error positivo o por infravaloración del rendimiento) da lugar a una "cautela" inicial que se demuestra en una menor aceleración del movimiento de subida. A tan estudiada situación se le denomina "efecto apalancamiento" o, por su terminología inglesa, "efecto *leverage*".

El modelo sería exactamente igual al anterior incorporando ahora esta variable ficticia "i" que toma valor 1 cuando la innovación del período previo fue negativa y cero en el resto de los casos

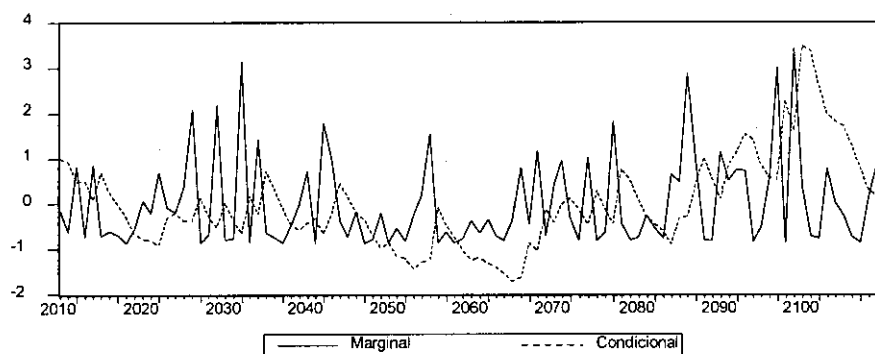
$$y_t = a_0 + a_1 h_{t-1} + \varepsilon_t$$

$$h_{t-1} = w + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2 i_{t-1} + \beta_1 h_{t-2}$$

Parámetros		Valor	Desviación Típica	Z- estadística	Prob.
Nombre					
Modelo de la varianza Condicional	OMEGA	0.047920	0.008763	5.468498	0.0000
	ALPHA(1)	0.058277	0.015174	3.840539	0.0001
	BETA	0.879790	0.016643	52.86253	0.0000
	ALPHA(2)	0.058279	0.017246	3.379189	0.0004
Modelo de la Media	MU	0.066060	0.047368	1.394610	0.0816
	THETA	0.120578	0.023675	5.093025	0.0000
	GAMMA	0.007697	0.036567	0.210482	0.4166
Log Verosimilitud		-3247.415	Akaike info criterion		3.090612
Media LogVerosimilitud		-1.541982	Schwarz criterion		3.109401
Nº de explicativas		7	Hannan-Quinn criter.		3.097493

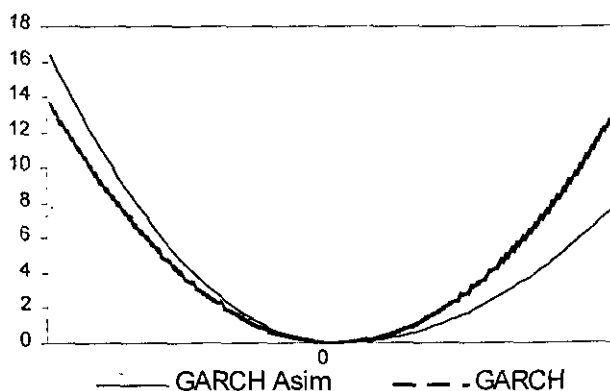
El contraste de Wald para la restricción del coeficiente  $\alpha_2$  igual a cero rechaza esta posibilidad con una probabilidad de más del 99%, hecho que ya se intuye a la vista de la z-estadística resultante en la salida presentada. En definitiva, parece plenamente justificada la introducción de una variable ficticia que recoja el carácter asimétrico de los movimientos de subida y de los de bajada.

Dado que la ficticia introducida solo modifica el valor de la estimación de la volatilidad cuando se producen infravaloraciones en el período anterior (en otro caso el parámetro estaría multiplicado por cero), los valores obtenidos para lo que incorpora el error previo en cualquier caso ( $\alpha_1 = 0,058277$ ) y lo que a esto se le añade en los momentos en los que el período anterior se infravaloró ( $\alpha_2 = 0,058279$ ) nos hace concluir que el efecto asimétrico es la duplicación en los momentos de infravaloración sobre los de sobrevaloración.



De cara a la comparación entre este y el modelo anterior, tiene sentido representar la Curva de Impacto de las Innovaciones (*News Impact Curve*) que definieran Engle y Ng (1993) y que quedó comentada en el correspondiente capítulo de contraste del modelo ARCH. Para la obtención de dicha curva se ha generado un programa de E-views cuyas instrucciones se encuentran en los anexos.

Curva de Impacto de las Innovaciones



En el gráfico, se ha representado la curva de impacto de las innovaciones tanto para los residuos de un modelo presentado como un GARCH (1,1) del tipo del modelo I como para un GARCH (1,1) con un factor asimétrico a partir de la entrada de la variable ficticia "i". La línea más gruesa corresponde al modelo con el GARCH asimétrico y la más delgada al GARCH normal.

Para el caso del GARCH normal, la pura matemática de la función empleada determina una representación parabólica totalmente simétrica. En el caso del GARCH asimétrico, la inclusión hace que el efecto de las innovaciones negativas aumente la volatilidad condicional estimada por encima del valor que le otorgaba el modelo GARCH (1,1); mientras que la volatilidad estimada tras períodos de "error positivo" el crecimiento de la volatilidad es más suave del que antes se enunciaba.

Este hecho también era evidente a partir de la comparación de los parámetros estimados. En el modelo I, el valor de  $\alpha_1$  era de 0,09; mientras que en el modelo II era de 0,05, a lo que habría que añadir el parámetro  $\alpha_2$  cuando el error previo fue negativo, luego en esos momentos se incorpora un producto por 0,10 frente al 0,9 del modelo I, y en los momentos de infravaloración del error previo, la volatilidad subsiguiente sólo se ve afectada por un producto de 0,05 en vez del 0,09.



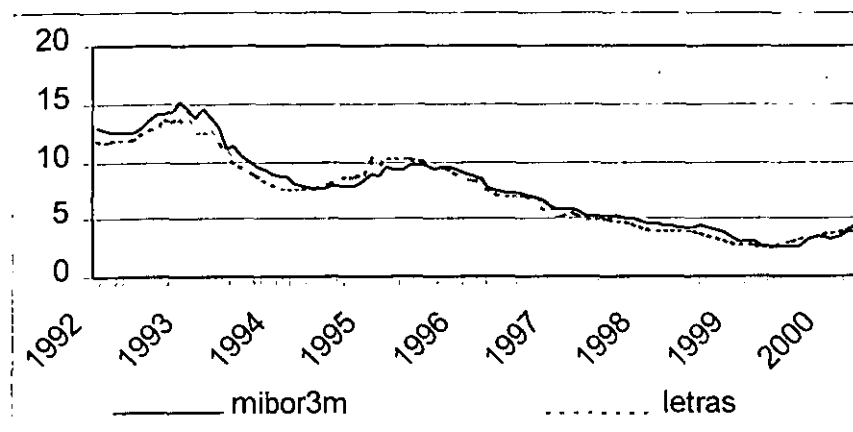
### **Modelo III: Modelo II activo sin riesgo en la explicación de la volatilidad - LNMIBOR3M -**

En diversas investigaciones sobre otros índices bursátiles se incorpora como explicativa de la volatilidad el activo sin riesgo, fruto evidente de la lógica de las teorías recogidas tanto en la CAPM de Sharpe (1964) y Litner (1965) como las APT de Ross (1976), como las formula de precios de opciones de Black y Scholes (1973) que han sido conveniente reflejadas en su correspondiente capítulo.

La disponibilidad estadística de las series de activos públicos no es de carácter diario en muchas ocasiones, sino media mensual, o bien no se han publicado convenientemente hasta hace relativamente poco tiempo, por lo que no contamos con suficientes datos. Por este hecho, diversos autores han tomado la determinación de incluir en la especificación diaria el mismo valor para todas las observaciones de un mismo mes.

En nuestra aplicación y para no restar valor al contenido informativo diferencial del día a día, hemos preferido incluir una "proxi" de la variable pública que representaría claramente el activo sin riesgo (probablemente la Letra del Tesoro a un año para el caso español). La variable elegida sobre la que sí hay disponibilidad diaria de datos es el tipo de interés interbancario a tres meses de Madrid (Mibor3m). Esta serie estadística recoge los tipo de interés medios diarios a los que se han cruzado operaciones de préstamo entre las entidades financieras con plaza de operaciones en Madrid. Dado que el tipo se mueve diariamente y que los activos públicos marcan el valor mínimo de los tipos de interés (por su condición de "sin riesgo") el seguimiento, con un cierto diferencial, del tipo de las letras por el MIBOR es casi perfecto, reflejando éste de forma casi inmediata los cambios en el otro, aunque quizá anticipando ligeramente las subidas y descontando con cierta tardanza las bajadas, pero con un margen de desfase temporal despreciable. En particular, para los años considerados en esta aplicación, el coeficiente de correlación entre el MIBOR a 3 meses y las letras del Tesoro a un año es de un 0,98. En la gráfica siguiente se puede observar como, durante todos estos años, el seguimiento es casi perfecto.

Tipos de interés MIBOR a 3 meses y Letras del Tesoro 1992-2000



En cuanto al plazo del tipo elegido, las disponibilidades estadísticas nos permitirían contar con el tipo de interés interbancario a un día, pero esta información incluiría condicionantes distintos a los que queremos reflejar con el activo sin riesgo, por cuanto incorporaría la problemática explícita y coyuntural del "cash flow" de la banca además de una cierta apuesta y reflejo sobre la rentabilidad de otros activos. El plazo de tres meses es el referente habitual, por lo que se convierte en un buen medidor de la evolución y de las expectativas sobre el tipo de interés en general y, por todo lo dicho, representa bien el papel que los tipos de interés puedan representar en la formación de la rentabilidad de su alternativa natural para el inversor; es decir, la rentabilidad bursátil.

En esta ocasión, el modelo del error propuesto para su maximización, se formula del siguiente modo:

$$\varepsilon_t = y - \mu - \theta \varepsilon_{t-1} - \gamma h_{t-1}$$

$$h_t = w + \alpha_1 \varepsilon_{t-1}^2 + \beta h_{t-1} + \alpha_2 \varepsilon_{t-1}^2 i_{t-1} + \pi \text{Ln} \text{mibor3m}$$

$$i = \begin{cases} \varepsilon_{t-1} < 0; 1 \\ \varepsilon_{t-1} > 0; 0 \end{cases}$$

La variable que se incorpora respecto al modelo anterior es  $\text{Ln} \text{mibor3m}$ , primera diferencia del logaritmo del valor diario del MIBOR a tres meses.

Los resultados de estimación iterativa máximo verosímil son los siguientes:

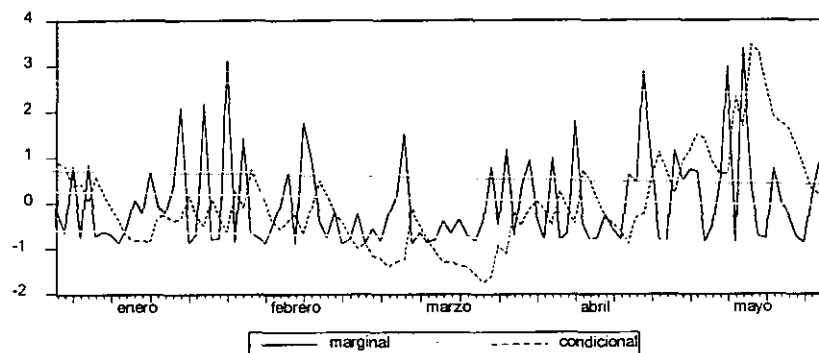
		Parámetros		Z-estadística	Prob.
	Nombre	Valor	Desv Típica		
<b>Modelo de la Varianza Condicional</b>	OMEGA	0.048154	0.008726	5.518694	0.0000
	ALPHA(1)	0.059906	0.015261	3.925341	0.0000
	BETA	0.879941	0.016437	53.53513	0.0000
	PI	1.059821	0.560919	1.889437	0.0294
	ALPHA(2)	0.052874	0.017850	2.962187	0.0015
<b>Modelo de la Media</b>	MU	0.065975	0.047892	1.377583	0.0842
	THETA	0.118384	0.023692	4.996734	0.0000
	GAMMA	0.007830	0.0036505	2.334494	0.0151
<b>Log Verosimilitud</b>		-3245.530	<b>Akaike info criterion</b>		3.08977
<b>Media LogVerosimilitud</b>		-1.541087	<b>Schwarz criterion</b>		3.11124
<b>Nº de explicativas</b>		8	<b>Hannan-Quinn criter.</b>		3.09763

- El valor de maximización del logaritmo de verosimilitud es de -3245, frente al anterior de -3247, con lo que se produce una ligera mejoría en el resultado obtenido. Los criterios de Akaike y de Schwarz coinciden en valorar positivamente la inclusión de una variable más, ya que se reducen ligeramente.
- La nueva variable incluida como explicativa es representativa con una significatividad de más del 98% y su introducción, no modifica de un modo importante el resto de los parámetros incluidos. El alto valor del parámetro estimado se debe a la unidad de medida de la primera diferencia del logaritmo del MIBOR a 3 meses que, en este caso, no ha sido multiplicada por cien como sí se ha hecho con la variable endógena por las razones que se comentaban en el modelo I.
- La suma de los parámetros del error y de la varianza condicional retardados es muy próxima a uno (0,97), manteniéndose el citado caso de "persistencia de la volatilidad que deberá ser objeto de algún tipo de corrección.
- Los resultados respecto al valor del parámetro del efecto apalancamiento siguen manteniendo el hecho ya apuntado en el modelo anterior de que la sobrevaloración del rendimiento en el período anterior produce un incremento

sobre la volatilidad condicional del período siguiente prácticamente igual al doble de cuando en el período anterior se produjo una infravaloración.

- Los valores de ajuste entre varianza marginal y condicional siguen mostrando el efectivo retardo entre ambas y, de cara al ajuste en los últimos momentos y a partir del gráfico del último año (de enero a junio del 2000), sigue presentándose un buen comportamiento del modelo para recoger los cambios de tendencia, salvando la "sorpresa" del mes de marzo, que aparece claramente mal representada por este modelo.

Gráfica de la varianza marginal y condicional estimada en el año 2000



En definitiva, la ganancia en este modelo respecto a los anteriores es evidente, por cuanto aunque no es enorme. La inclusión de la variable de activo sin riesgo es a todas luces un factor explicativo a considerar por su lógica financiera y por su significatividad estadística.

#### **Modelo IV: Modelo III con corrección estacional de los residuos**

Tal y como se ponía de manifiesto en el capítulo referido a los modelos GJR-ARCH y como viene siendo práctica generalizada en la modelización de las variables de tipo financiero (y, en particular, en las de rendimientos bursátiles), es importante la consideración del mes y del día para determinar la mayor o menor volatilidad. En general, parece contrastado que períodos de dilación en la actividad dan lugar a mayores signos de volatilidad cuando se reinicia la actividad normal. El caso del lunes, día que evidentemente es el primero después de dos de inactividad bursátil, o períodos como septiembre y octubre, seguidos del "descanso vacacional", se muestran enormemente volátiles en comparación con otros meses o días del año.

La aportación más original del modelo GJR-ARCH se refiere a la incorporación de la perturbación desestacionalizada en la estimación de la varianza condicional. Efectivamente, es común en los análisis de tipo financiero hacer referencia a volatilidades con una particular estacionalidad, tanto en la frecuencia mensual, como en la frecuencia semanal.

Para comprobar las posibles diferencias debidas al día de la semana o al mes al que se refiera el dato, se ha calculado la volatilidad media (varianza) en el rendimiento sobre el día anterior tanto para cada día de la semana como para mes del año. Los resultados, que aparecen en las gráficas posteriores, se resumen en lo siguiente:

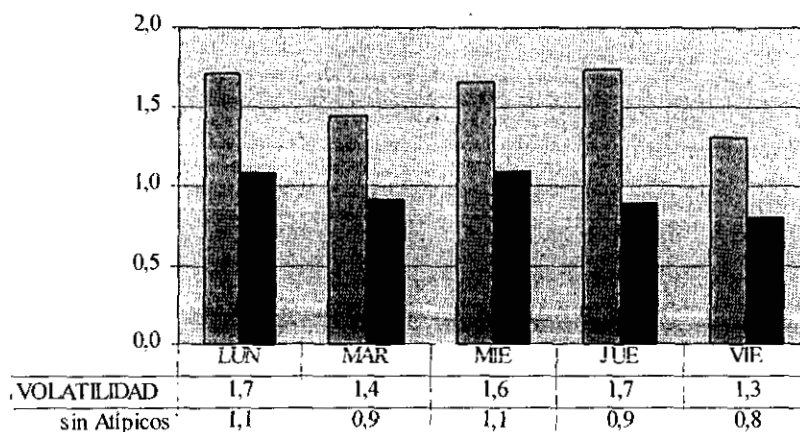
- en el espacio de los cinco días hábiles normales de la semana (de lunes a viernes, ambos inclusive, en la bolsa de Madrid), los valores de la varianza más significativos se producen los lunes y los jueves, aunque las diferencias con el resto de los días son apenas imperceptibles. Más aún, cuando se eliminan de la serie de cálculo todos aquellos días que han tenido un crecimiento especialmente alejado del comportamiento normal (atípicos)<sup>28</sup>, las diferencias entre los diferentes días de la semana aún son menos perceptibles, aunque sí es verdad que el lunes sigue figurando como el día más volátil. Diversos analistas dicen de esta realidad

---

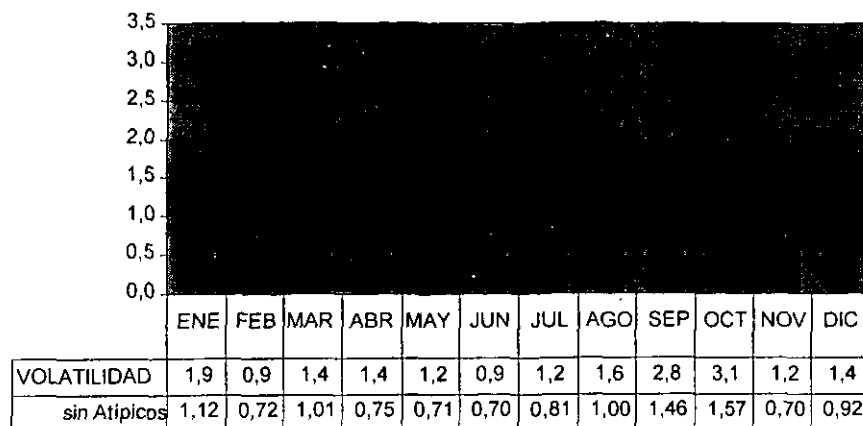
<sup>28</sup> De cara a este análisis, se han considerado atípicos todos aquellos puntos de crecimiento inter-día que estuvieran por encima o por debajo de  $\pm 2$  veces la desviación típica de la serie de crecimientos correspondientes a ese día de la semana o a ese mes.

empírica que no es más que un reflejo lógico del día que, en todos los casos, tiene un período de maduración de los activos no inferior a los tres días (desde, por lo menos, el viernes, hasta el lunes, con dos días intermedios sin actividad bursátil en Madrid).

Evaluación de la varianza en los distintos días de la semana (enero 1992 a junio 2000)



- Para el caso de los meses (realizando nuevamente el estudio anterior sobre varianzas y atípicos) las diferencias en la volatilidad de cada mes sí son mucho más representativas. Tanto en términos absolutos, como descontados los atípicos habidos durante estos diez años, parece clara la observación de que los meses de septiembre y octubre son los más volátiles del año. Comparando con los valores previos y posteriores, también es importante la volatilidad registrada en el mes de enero respecto a diciembre o a febrero, aunque, en ningún caso, la volatilidad de este mes es tan grande como la observada en los dos casos anteriormente descritos.

Evaluación de la varianza en los distintos meses (enero 1992 a junio 2000)

En este modelo IV la propuesta que se realiza parte de un filtrado previo de los datos a partir de las variables ficticias estacionales, eliminando del fenómeno a modelizar este componente debido al mes que ya se puede añadir de forma automática. La especificación matemática elegida sería la siguiente:

$$\begin{aligned}\varepsilon_t &= y_t - \mu - \theta\varepsilon_{t-1} - \gamma h_{t-1} \\ \eta_t &= \varepsilon_t (1 + \lambda_2 \text{enero} + \lambda_3 \text{diciembre}) \\ h_t &= w + \alpha_1 \eta_{t-1}^2 + \beta h_{t-1} + \alpha_2 \eta_{t-1}^2 i_{t-1} + \pi \text{LNMIBOR3M}_t\end{aligned}$$

donde:

- Enero y diciembre son dos variables ficticias que toman valor uno para observaciones incluidas en el mes que representan y cero en el resto de los casos. La razón de incluir estos meses y no otros es fruto de una serie de regresiones realizadas para comprobar cuáles eran los meses estadísticamente significativos.
- Lnmibor3m es la primera diferencia del logaritmo del valor diario del MIBOR a tres meses.

Siguiendo los mismos pasos de Engle y Ng (1993), para la selección de las variables ficticias a emplear se ha practicado una regresión inicial del valor del residuo del modelo de medias móviles de la variable objetivo con una ficticia para cada mes (el abultado

volumen de los grados de libertad disponibles con un muestra de más de 2100 observaciones nos permite despreciar la pérdida por inclusión de más variables explicativas a la hora de calificar la t-estadística obtenida). Las variables a priori significativas para la estimación de dicho error al cuadrado resultaron ser las de enero y diciembre.

	Parámetros		Desv Típica	Z- estadística	Prob.
	Nombre	Valor			
<b>Modelo de la Varianza Condicional</b>	OMEGA(1)	0.048466	0.008957	5.411194	0.0000
	ALPHA(1)	0.054893	0.014578	3.765457	0.0001
	BETA(1)	0.880695	0.016399	53.70456	0.0000
	PI(1)	1.084190	0.560946	1.932788	0.0266
	ALPHA(2)	0.060405	0.018045	3.347455	0.0004
<b>Modelo de la Media</b>	MU(1)	0.047914	0.048856	0.980722	0.1634
	THETA(1)	0.118502	0.023932	4.951509	0.0000
	GAMMA(1)	0.007289	0.003654	2.994844	0.4209
<b>Corrección estacionalidad</b>	LAMBDA(2)	0.079013	0.028253	2.796541	0.0129
	LAMBDA(3)	0.180688	0.089822	2.011619	0.0221
<b>Log Verosimilitud</b>		-3243.881	<b>Akaike info criterion</b>		3.09010
<b>Media Log Verosimilitud</b>		-1.540304	<b>Schwarz criterion</b>		3.11694
<b>Nº de explicativas</b>		10	<b>Hannan-Quinn criter.</b>		3.09993

Sobre estos resultados, comentamos los siguientes puntos:

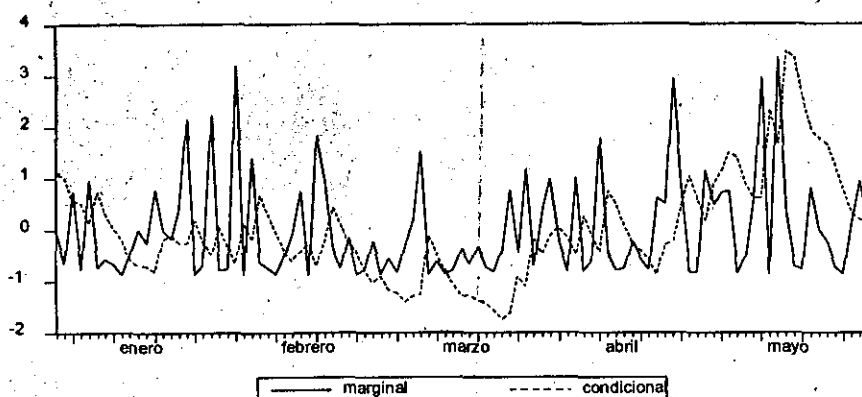
- Los parámetros de pre-filtrado estacional son significativos a un nivel superior al 98% y, por su valor positivo, incorporan mayor volatilidad condicional en el modelo en aquellos momentos en los que cobran efectividad, es decir en el mes que marcan.
- Nuevamente asistimos a una ganancia limitada, pero positiva, en el valor del logaritmo de verosimilitud maximizado, hecho que en esta ocasión no traslucen los estadísticos de Akaike o de Schwarz, que permanecen prácticamente en los mismos niveles que obteníamos en el modelo anterior, pero, en cualquier caso, no hablan de una pérdida de grados de libertad no compensada por reducción del error, sino que se mantienen en los mismos niveles previos.
- Por primera vez, el parámetro que recoge el efecto apalancamiento ( $\alpha_2$ ) es ligeramente superior al del impacto normal ( $\alpha_1$ ); es decir, el modelo propuesto



enfatisa aún más los incrementos en la varianza ante infravaloraciones en el período precedente, con lo que se acentúa la pendiente en "la caída". Para el comportamiento en períodos de sobrevaloración previa, el impacto sobre la volatilidad se mantiene en los términos que veníamos arrastrando desde el primer modelo presentado ( $\alpha_1=0,054893$ ).

- La primera valoración de predicción realizable a partir de la contrastación gráfica nos permite observar el mismo hecho ya comentado para otros modelos previos de adelanto de la varianza marginal sobre la condicional con ese efecto adicional de valoración del grado de acierto anterior. Por otro lado, el desfase en marzo de 2000 que ya se había comentado anteriormente sigue presente en esta especificación.

#### Varianza marginal y varianza condicional estimada en el año 2000



Este modelo IV supone, a todas luces, una ganancia en la especificación planteada respecto a los anteriores y viene a confirmar el hecho teórico de la incidencia especial de algunos períodos estacionales en la formación de la volatilidad.

En el reflejo de la varianza observada durante la década de los noventa que se llevaba a cabo al principio de este modelo IV, claramente eran los meses de septiembre y octubre aquellos que históricamente habían mostrado valores más grandes. En multitud de textos financieros se refiere este hecho y se pone en relación con el momento de reemprender al actividad después de la pausa del verano. Sin embargo, es claramente contrastable que cada vez se produce menos esa "pausa", cuando menos en términos tan contundentes

como cuando prácticamente se suspendía la actividad. Sin duda los meses de "inicio de curso" son importantes en la generación de expectativas y son la preparación de los análisis sobre los resultados, aunque estos no se conocerán realmente hasta finales de año. También octubre es el momento en que el gobierno perfila sus presupuestos para el año siguiente. Diversas crisis financieras han marcado días de octubre como días negros confirmando este hecho.

A pesar de todo lo dicho, el modelo no muestra estos meses como básicos en el comportamiento de la volatilidad y sí los meses de diciembre y enero que, en principio presentan valores algo menos volátiles que septiembre y octubre. Evidentemente estos meses marcan dos elementos que han de afectar a la volatilidad:

- El anuncio de los resultados sociales y de los dividendos
- La entrada en vigor de diversas disposiciones jurídicas entre ellas económicas (nuevo presupuesto, modificaciones fiscales, etc.).

Aún así, cabe plantear la duda de si lo que están recogiendo estas variables ficticias es verdaderamente un comportamiento estacional de mayor volatilidad condicional trasladable a futuro o una serie de coincidencias históricas en estos meses que hacen que estas ficticias no sean otra cosa que una corrección sobre el modelo inicial, sin que tenga porqué repetirse en el futuro.

Durante la década, diversos acontecimientos de inmensa trascendencia económica han fijado su fecha de inicio durante los meses citados:

- Diciembre 1992 - enero de 1993. Crisis del Sistema Monetario Europeo que deparó la salida de la Lira italiana y de la libra esterlina y diversas devaluaciones de la peseta.
- Enero de 1993: Entrada en vigor del Mercado Único europeo y la libertad de movimiento de capitales en España.
- Diciembre 1996 - enero 1997: recrudecimiento de la crisis asiática con la no paridad del yuan en Indonesia.

- Diciembre 1997 - enero 1998: crisis cambiaria latinoamericana.
- Enero de 1999- entrada en funcionamiento del Banco Central Europeo y creación del Euro
- Enero de 2000: creación de los Índices tecnológicos y explosión de este tipo de activos en cuanto a récord de subidas y bajadas súbitas.

Esta coincidencia histórica nos debe llevar a ser cautelosos con el carácter de estas variables y su correcta interpretación.

Aún así, el modelo GJR-ARCH con filtrado estacional sobre los meses de enero y diciembre parece ser el mejor de cuantos hemos planteado hasta el momento, tanto por ganancia en el valor maximizado como por la significatividad de todas las variables incluidas en dicho modelo.

### Modelo IV alternativo: GJR-ARCH + ARCH de memoria larga (Modelo IV + ARCH de componentes)

Los resultados obtenidos para el modelo que hemos llamado IV reflejan una importante "persistencia en volatilidad" en el sentido en que definían esta Engle y Bollerslev (1986). La suma paramétrica es de  $\alpha_1 + \alpha_2 + \beta_1 = 0,9658$ . Atendiendo a la medida aproximada de Glosten y otros (1993) que ya comentábamos anteriormente, el autorregresivo de orden uno de la varianza condicional estimada tiene un parámetro de 0,87, significativo con más del 99% de probabilidad. En definitiva, parece necesario llevar a cabo algún tipo de estimación que corrija el efecto de integración de las varianzas condicionales y es por ello que se plantea un modelo compuesto del anterior (GJR-ARCH) y de los conocidos recientemente como de memoria larga o ARCH de componentes.

Tal y como deducían Ding y Granger (1997) su modelo de memoria larga<sup>29</sup> se puede escribir como un GARCH(2,2) con parámetros restringidos de modo que la varianza condicional se explicitaría como:

$$h_t = \sigma^2(1-w)\alpha_1(1-\alpha_2-\beta_2) + [w\alpha_1 + (1-w)\alpha_2]\varepsilon_{t-1}^2 - [w\alpha_1\beta_2 + (1-w)(1-\alpha_1)\alpha_2]\varepsilon_{t-2}^2 + (1-\alpha_1-\beta_2)h_{t-1} - (1-\alpha_1)\beta_2h_{t-2}$$

donde  $\sigma^2$  responde a la varianza marginal del proceso. Siguiendo este esquema se ha especificado el modelo GJR-ARCH que ya se había planteado añadiendo ahora estas dos componentes a la varianza. En detalle, la estimación se realiza sobre el siguiente planteamiento:

$$\begin{aligned}\varepsilon_t &= y - \mu - \theta\varepsilon_{t-1} - \gamma h_{t-1} \\ \eta_t &= \varepsilon_t(1 + \lambda_2\text{enero} + \lambda_3\text{diciembre}) \\ h_t &= (1-\alpha-\beta)w + (\alpha_1 + \phi_1)\eta_{t-1}^2 - (\alpha_1 + \beta_1)\phi_1\eta_{t-2}^2 + \\ &(\beta - \phi_1)h_{t-1} - (\beta\rho - (\alpha + \beta)\phi_1)h_{t-2} + \alpha_2\eta_{t-1}^2i_{t-1} + \pi\text{LNMBOR3M}_t\end{aligned}$$

En definitiva, a la especificación de la varianza mediante un activo sin riesgo (representado por  $\text{Lnmbor3m}$ ), un efecto asimétrico (marcado por la "i"), a la función

<sup>29</sup> Ver formulación de los modelos IGARCH

GARCH se le incorpora un efecto a corto plazo, con sus correspondientes posibles intervinientes, y un efecto a largo plazo.

Los resultados obtenidos para los parámetros de este nuevo modelo son comparados con los anteriores en el cuadro siguiente. De cara a la convergencia matemática del modelo ha sido necesario realizar la estimación eliminando en este caso el primer año completo de la muestra (1992). De cara a hacer comparables los resultados, abajo se exponen los parámetros del modelo anterior también estimados para ese periodo.

Comparación de resultados entre los modelos IV y V

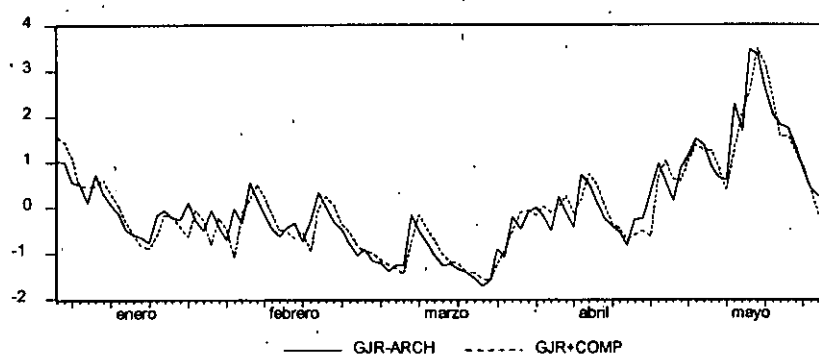
	GJR ARCH		GJR +COMPONENT ARCH		Diferencia %
	Parámetro	z-estadística	Parámetro	z-estadística	
OMEGA	0,049	5,160	-4,340	-4,152	-101,1%
ALPHA	0,047	3,367	0,054	1,882	-12,5%
BETA(1)	0,889	54,677	0,958	2,953	-7,3%
ALPHA(2)	0,009	1,413	0,072	2,838	-87,0%
CHI	0,059	3,258	0,011	1,587	437,9%
MU	0,056	1,037	0,073	1,383	-23,2%
THETA	0,118	4,731	0,129	5,866	-8,3%
GAMMA	0,015	0,382	-0,005	-0,117	-431,0%
LAMBDA(2)	0,086	1,842	0,092	1,920	-6,0%
LAMBDA(3)	0,158	2,705	0,128	1,410	23,6%
PHI	-	-	-0,073	-1,747	-
RHO	-	-	0,084	0,332	-
Log máxima verosimilitud	-2.959	-	-2.956	-	0,1%
Media Log máxima verosimilitud	-1.5522	-	-1.5508	-	-
Número de parámetros	10	-	12	-	-

Al margen de los términos constantes, donde la inclusión de nuevas variables está detrás de los importantísimos cambios registrados, el modelo planteado presenta las siguientes características comparadas:

- las variables nuevas incluidas no son significativas para la determinación de la volatilidad (a juzgar por los estadísticos t-student calculados, no significativos ni para el parámetro rho ni para el parámetro phi).

- En el nuevo modelo dejan de ser relevantes tanto los valores de corrección estacional del error, como el rendimiento del activo sin riesgo y, más importante, la varianza condicional como explicativa del modelo en media.
- En general, los parámetros calculados van casi hasta la nulidad en todos los casos, con una importantísima reducción en el efecto asimétrico, por ejemplo.
- Unido a todo lo anterior, el modelo supone una ganancia en cuanto al valor maximizado del logaritmo de verosimilitud de apenas un 0,1%.
- Comparando gráficamente los resultados obtenidos para la volatilidad estimada con el modelo IV y este GJR-Component ARCH, las diferencias de estimación son muy pequeñas, aunque la nueva estimación parece responder a un valor aún más desplazado hacia la derecha de la volatilidad condicional estimada frente a la varianza marginal. Obviamente esta situación responde a la "memoria prolongada" que subyace en el modelo ARCH de componentes.

Gráfica comparativa modelo GJR-ARCH y GJR-ARCH con componentes



A la luz de lo expuesto en los puntos anteriores, no parece conveniente la introducción de este modelo de memoria larga a pesar de los valores de que la suma de los parámetros estimados anteriormente fueran próximos al valor que haría al modelo no estacionario. Por supuesto, sí tiene sentido realizar un esquema tipo modelo exponencial como el E-GARCH propuesto por Nelson (1991).

### Modelo V: GJR-ARCH con estructura E-GARCH (Modelo IV con errores normalizados y en logaritmos)

Como ya quedaba extensamente expuesto en su capítulo correspondiente, Nelson (1991) alertaba sobre algunas situaciones cuando menos "comprometidas" en el empleo de los modelos ARCH. Algunas de ellas, como el efecto asimétrico, ya han sido debidamente valoradas en el modelo que venimos proponiendo pero otras no menos importantes como la posibilidad de que el resultado de la estimación de la varianza condicional pudiera salir un valor negativo siguen siendo un potencial problema.

En nuestro caso, la inclusión de la variable del activo sin riesgo puede ocasionar dicho efecto no deseable y ello a pesar de que todos los parámetros son positivos. En la medida en que el MIBOR (representado su rendimiento por la diferencia de su logaritmo; al igual que la endógena) sea suficientemente negativo podría darse el caso de un resultado total de la varianza condicional negativa.

La propuesta de este sexto modelo sería incluir una estructura exponencial del tipo Nelson (1991) cuya formulación concreta sería la siguiente:

$$\varepsilon_t = y_t - \mu - \theta \varepsilon_{t-1} - \gamma h_{t-1}$$

$$\eta_t = \varepsilon_t (1 - \lambda_2 \text{enero} - \lambda_3 \text{diciembre})$$

$$\log(h_t^{1/2}) = \log(w) + \alpha_1 \eta_{t-1}^2 / \sqrt{h_{t-1}} + \beta \log(h_{t-1}) + \pi \ln \text{mibor3m} + \alpha_2 \eta_{t-1}^2 i / \sqrt{h_{t-1}}$$

$$i = \begin{cases} \varepsilon_{t-1} < 0; 1 \\ \varepsilon_{t-1} > 0; 0 \end{cases}$$

Los resultados de la regresión planteada, unión del modelo GJR-ARCH y del planteamiento EGARCH de Nelson, se resumen en la siguiente tabla de información:

	Parámetros		Desv. Típica	z- estadística	Prob.
	Nombre	Valor			
<b>Modelo Varianza Condicional</b>	OMEGA	0.950502	0.005914	160.7223	0.0000
	ALPHA(1)	0.041720	0.009129	4.570318	0.0000
	BETA	0.932381	0.009744	95.68895	0.0000
	PI	0.815553	0.380010	2.146133	0.0159
	ALPHA(2)	0.028687	0.010945	2.621097	0.0044
<b>Modelo de la Media</b>	MU	0.048362	0.049634	0.974376	0.1649
	THETA	0.117807	0.023562	4.999853	0.0000
	GAMMA	0.010875	0.036491	2.298012	0.3828
<b>Corrección estacionalidad</b>	LAMBDA(2)	0.086759	0.097193	1.892645	0.0510
	LAMBDA(3)	0.148160	0.090940	1.629201	0.0516
<b>Log Verosimilitud</b>		-3247.445	<b>Akaike info criterion</b>		3.09349
<b>Media LogVerosimilitud</b>		-1.541997	<b>Schwarz criterion</b>		3.12033
<b>Nº de explicativas</b>		10	<b>Hannan-Quinn criter.</b>		3.10331

Como cabía esperar, la estimación tipo E-GARCH produce algunos cambios en los resultados obtenidos, de los que se podrían destacar los siguientes:

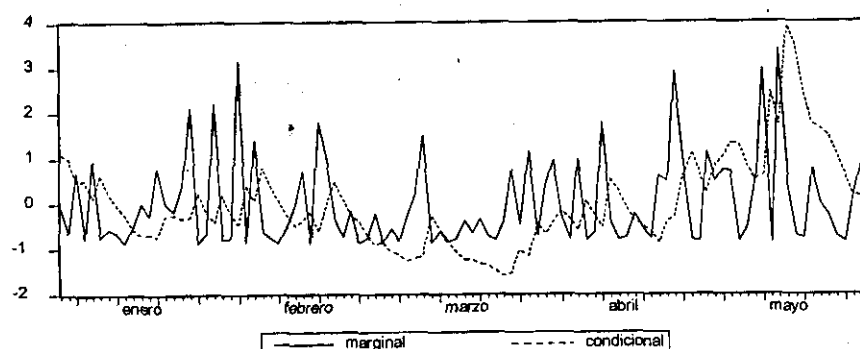
- El parámetro que representa el efecto apalancamiento se reduce considerablemente con relación al que afecta a la innovación cuadrada en el momento anterior del tiempo (0,02 y 0,04, respectivamente) pero la estructura logarítmica del modelo planteado no nos permite hablar de que un valor sería la mitad que el otro. Si aproximamos la diferencia entre los parámetros a sus valores exponenciales para hacer a ésta comparable con las que hemos realizado anteriormente, dicha diferencia sería casi nula (menos de 0.01), con lo que los valores estimados serían similares a los que obteníamos en modelos anteriormente planteados.
- Los parámetros de corrección de estacionalidad del modelo dejan de ser significativos para valores del 95%, cuando menos claramente el que afecta al mes de enero. Este hecho contradice claramente lo que se ha venido observando en el resto de las especificaciones así como lo que apunta la lógica financiera.

Algunas circunstancias nos hacen cuestionar la aplicación en este caso de un modelo EGARCH:



- La variable original sobre la que modelizamos el valor del error al cuadrado para la obtención de la varianza condicional, ya responde a una estructura logarítmica (es la diferencia de orden uno del logaritmo del precio diario del Ibex 35). Por esta razón, los residuos obtenidos también comportan una estructura logarítmica y, con el modelo EGARCH, estaríamos, en cierto modo, modelizando el logaritmo del logaritmo, con todo lo que ello pueda tener de complicación matemática y, sobre todo, de reducción de los valores hasta cifras tan pequeñas que la dificultad para la convergencia de los algoritmos de resolución se ve ampliamente trastocada.
- En segundo lugar, la estructura de los errores y su grado de ajuste frente a la serie del error al cuadrado encuentra muchos más puntos de error en cuanto a determinación de una tendencia totalmente inversa a la observada en la varianza marginal, hecho, en principio, poco probable.

Gráfica de la varianza marginal y condicional estimada para el año 2000



- En tercer lugar, comparando las medidas de error propuestas como contrastes sobre la capacidad predictiva, el modelo GJR-ARCH es claramente mejor que el modelo GJR - EGARCH.

MEDIDA DE ERROR	GJR - ARCH	GJR - EGARCH
MAE	1.546042	1.8611173
LMAE	1.820851	4.0571025
MSE	10.660236	104.10747
LMSE	7.090483	37.308427

- El efecto asimétrico incorporado por el modelo EGARCH, importante aportación de su creador, ya ha sido incluido en el modelo anterior con la variable ficticia "i", con lo que, sobre este particular, no hay ninguna novedad en esta especificación, que, eso sí, sigue confirmando la significatividad de la variable introducida.
- Finalmente, al aplicar el modelo EGARCH el logaritmo de verosimilitud obtenido es de -3247 frente al del modelo GJR-ARCH anterior de -3243, con lo que la ligera pérdida es clara en cuanto a la potencia del modelo propuesto por Nelson (1991) para la serie a la que aquí se está aplicando.

En definitiva, no hay una ganancia clara que justifique sin lugar a dudas la utilización de un modelo E-GARCH frente al que planteábamos como modelo IV o modelo GJR-ARCH, y, sin embargo, diversas apreciaciones más pragmáticas que teóricas, hacen más plausible el empleo del modelo tipo GJR.

### **Comparación conjunta de los modelos I al V**

Con el fin de poner de manifiesto todos los elementos tomados en cuenta para decidir que el modelo IV es, hasta el momento, el que mejores resultados presenta en la modelización tipo ARCH, se presentan en un cuadro agregado los contrastes sobre su bondad de ajuste y estructura tal y como se especificaban en la correspondiente sección del esquema teórico.

## Comparación entre los modelos y ganancias obtenidas

	Modelo I		Modelo II		Modelo III		Modelo IV		Modelo V	
	Valor	Contraste	Valor	Contraste	Valor	Contraste	Valor	Contraste	Valor	Contraste
Asimetría	0,056		-0,274		-0,250		-0,250		-0,159	
Curtosis	2,293		6,166		6,150		6,149		4,625	
Jarque Bera	15,295	34%	906,229	0%	893,651	0%	891,786	0%	24,671	0%
Test sesgo signo	-1,435	-2,850	-0,050	-1,170	0,057	1,317	-0,061	-1,398	-1,440	-0,060
sesgo negativo	0,054	1,920	0,030	0,843	0,020	0,820	0,030	0,830	-0,026	-0,725
sesgo positivo	-0,043	1,270	0,009	0,259	0,007	0,004	-0,110	-1,240	0,000	-0,010
Sesgo Conjunto	F =	0,98	F =	0,64	F =	0,98	F =	1,06	F =	1,32
Logaritmo de máxima verosimilitud	-3252,05		-3247,42		-3245,53		-3243,88		-3247,45	
Mae	1,563		1,601		1,548		1,546		1,598	
Lmae	1,821		1,903		1,822		1,820		1,829	
Mse	10,908		11,986		10,687		10,620		11,448	
Lmse	7,048		7,404		7,095		7,090		7,183	
Persistencia en volatilidad (alpha+beta)	0,972		0,996		0,993		0,993			
AR(1) de ht	0,874	17,95	0,878	18,57	0,972	189,15	0,971	186,93	0,879	18,76
Evolución de los parámetros										
Modelo I			Modelo II			Modelo III			Modelo IV	
OMEGA(1)	0,0434	0,0479	0,0482	0,0485						
ALPHA(1)	0,0939	0,0583	0,0599	0,0549						
BETA(1)	0,8781	0,8798	0,8799	0,8807						
ALPHA(2)	-	0,0583	0,0529	0,0604						
PI(1)	-	-	1,0598	1,0842						
MU(1)	0,0736	0,0661	0,0660	0,0479						
THETA(1)	0,1187	0,1206	0,1184	0,1185						
GAMMA(1)	0,0153	0,0077	0,0078	0,0073						
LAMBDA(2)	-	-	-	0,0790						
LAMBDA(3)	-	-	-	0,1807						
Parámetros estandarizados										
Modelo I			Modelo II			Modelo III			Modelo IV	
ALPHA(1)			0,2668	0,1649	0,1696	0,1553				
BETA(1)			0,9537	0,9796	0,9635	0,9658				
ALPHA(2)			-	0,1649	0,1497	0,1709				
PI(1)			-	-	0,0117	0,0120				
Desv típica del error			3,59019	3,5774	3,5795	3,5768				
Desv típica volatilidad estimada			1,3728	1,4075	1,3841	1,3862				
Desv típica de LNPRECIO*100			1,26409	1,26409	1,2641	1,2641				
Desv típica de LNmbor3m			-	-	0,0140	0,0140				

## **No normalidad**

El carácter asimétrico de los residuos es evidentemente negativo para todos los modelos planteados salvo para el primero, aunque los valores del coeficiente de asimetría no están muy lejanos de cero. El valor que definitivamente excluye un supuesto de normalidad, por lo menos en la representación muestral del residuo estimado, es el valor de curtosis, que en cualquiera de los casos, está muy lejos del valor tres que sería el esperado. El contraste Jarque Bera no deja lugar a dudas sobre este hecho, siendo tan solo el primer modelo el más cercano a una distribución normal aunque con una probabilidad escasísimamente significativa (34%). Todos los demás modelos tienen probabilidad cero de ser una normal a juzgar por los valores de este contraste.

Hay que decir que la incorporación de especificaciones más complejas que han ido dando lugar a los cinco modelos presentados no nos ha aproximado en absoluto a la distribución de probabilidad normal que se supone teóricamente y con la cual se han ido estimando los parámetros. Sin embargo, quedaba patente en la primera modelización propuesta que la estimación en función de distribuciones probabíísticas diferentes (la t-student o la Distribución Generalizada de los Errores - GED - ) no suponía ninguna mejora en los resultados obtenidos.

## **Contrastes de asimetría (sesgo de signo)**

Los contrastes de sesgo de signo para el primer modelo muestran un carácter claramente significativo del valor asimétrico del residuo (la t experimental es tiene un valor de 2,8). Dicho sesgo, a juzgar por los contrastes parciales de signo negativo y positivo, parece estar centrada en el primero de estos. La inclusión de la variable ficticia de discriminación en función del signo del residuo en el período anterior corrige perfectamente este sesgo, pudiendo observarse que para el resto de los modelos planteados, los contrastes de sesgo de signo son no significativos en todos los casos.

## **Logaritmo de máxima-verosimilitud**

La evolución del logaritmo de verosimilitud ha sido siempre mejor a medida que se han ido incluyendo modelizaciones más complejas. Siempre con valores negativos, la ganancia (valor mayor) ha sido una constante en cada modelo propuesto, salvo en el último, en el que se proponía la modelización tipo E-GARCH. Hay que destacar que

los valores de ganancia en el logaritmo pueden parecer algo reducidos, pero no hay que olvidar que se está modelizando una variable diferencia de logaritmos, donde la escala es bajísima.

### Ratios del error muestral

Para determinar diferencias significativas entre los diversos modelos ARCH posibles, algunos autores recogen distintas funciones de pérdida que evalúan la gravedad del error cometido. De entre ellos, los más comunes son los que desarrollaran Hamilton y Susmel (1994) y que se resumen en las siguientes cuatro propuestas:

#### Expresión de los contrastes sobre el error

	<i>Directo</i>	<i>Valores Absolutos</i>
<b>Errores</b>	$MSE = \frac{\sum_{t=1}^T (\varepsilon_t^2 - h_t)^2}{T}$	$MAE = \frac{\sum_{t=1}^T  (\varepsilon_t^2 - h_t) }{T}$
<b>Logaritmo de errores</b>	$LMSE = \frac{\sum_{t=1}^T (\ln(\varepsilon_t^2) - \ln(h_t))^2}{T}$	$LMAE = \frac{\sum_{t=1}^T  (\ln(\varepsilon_t^2) - \ln(h_t)) }{T}$

Donde  $h_t$  es la varianza condicional de  $\varepsilon_t^2$

Observando estos valores para los modelos propuestos, el segundo modelo parece presentar un ligero retroceso respecto al primero; pero los dos siguientes son claramente mejores que el primero y que el segundo. Todos los contrastes siguen una evolución decreciente (menos error) en los modelos uno tres y cuatro. El modelo V (estructura E-GARCH) nuevamente supone peores resultados respecto a los anteriores y, en este caso, incluso respecto al modelo I.

Valores representativos del error muestral

Modelo	I	II	III	IV	V
Mae	1,563	1,601	1,548	<b>1,546</b>	1,598
Lmae	1,821	1,903	1,822	<b>1,820</b>	1,829
Mse	10,908	11,986	10,687	<b>10,620</b>	11,448
Lmse	<b>7,048</b>	7,404	7,095	7,090	7,183

(En negrita, los mejores de cada contraste)

El único valor ciertamente particular de la tabla anterior se refiere al mínimo error en valores logarítmicos de la suma media de los errores al cuadrado, donde el modelo I presenta el mejor resultado, aunque no muy distante del valor del modelo IV elegido y, probablemente, debido a un punto singular casualmente mejor recogido por el primero de los modelos propuestos.

Si se puede considerar a esta medida algo significativo del ajuste, es interesante poner de manifiesto que el modelo E-GARCH presenta peores resultados en todos los contrastes incluso ya para el modelo III presentado.

En definitiva, salvo algún elemento puntual, las funciones de pérdida apuntan hacia el modelo cuarto como el mejor de los propuestos.

### Persistencia en la volatilidad

Desde el primero de los modelos tipo ARCH presentados, la presencia de un elevado valor de los parámetros del proceso autorregresivo subyacente nos avisaban sobre el problema de la integrabilidad del sistema con el que estamos trabajando, aunque, en puridad de términos, la estimación estaba asegurada ya que los coeficientes sumados, si bien están muy próximos a uno, son siempre estrictamente menores.

Ateniéndonos a la medida de "persistencia en volatilidad" propuesta por Glosten y otros (1993), podemos comparar el parámetro de un autorregresivo de orden uno calculado con la serie muestral de la varianza condicional regresada sobre si misma en el momento anterior. Para todos los modelos propuestos, el valor del parámetro es cercano a uno y significativamente distinto de cero para más del 95% de probabilidad.

Llama la atención que el valor de los parámetros del contraste tipo Glosten sean muy similares en los modelos I y II estimados (0,87) y que se torne casi uno para los modelos III y IV, corrigiéndose nuevamente con la utilización del modelo tipo E-GARCH. Quizá esta es una de las únicas ventajas claras que presentaría este tipo de modelización que hemos rechazado previamente: eliminaría el efecto de fortísima memoria a largo plazo del modelo con el que estamos trabajando.

Como se exponía en la sección de esta tesis dedicada a los modelos GARCH integrados, una solución al problema que necesariamente hay que intentar es probar un modelo de memoria larga o ARCH de componentes, y dicha modelización alternativa es la que recogíamos con el modelo IV alternativo sin que los resultados aconsejaran su empleo.

### **Diferencias en los parámetros y valores estandarizados**

La evolución respecto al valor representativo del valor medio de la varianza marginal representado por el término constante ( $\omega$ ) no sufre apenas cambios en cualquiera de los cuatro modelos anidados que se han ido presentado. Este valor estaría en torno a 0.047.

Tampoco el parámetro representativo del valor autorregresivo de la varianza condicional ( $\beta$ ) cambia en exceso a pesar de la inclusión de más variables o de la corrección estacional. Dicho valor se centra en 0.87 para los cuatro modelos propuestos, con ligeros cambios.

El carácter autorregresivo del primer modelo ( $\alpha_1$ ) viene a representar la suma de los dos parámetros en los que se divide el efecto en el resto del modelo, por lo que el cambio tampoco es enorme en este caso, aunque sí claramente diferenciador cuando se incluye la variable discriminadora sobre el signo del residuo observado en el período precedente.

Es muy interesante señalar cómo ha basculado la importancia del efecto negativo en términos relativos al positivo en las modelizaciones presentadas. En el modelo II (primero con dicho término), la importancia del efecto apalancamiento se encuadra en valores muy similares, sino exactos, a los del autorregresivo normal sobre el valor de los residuos al cuadrado. En el modelo III, este efecto diferencial es, en valor absoluto

del parámetro, algo menor; y, en el último modelo propuesto, es superior al caso normal. Aún así, las diferencias en todos los modelos propuestos son muy pequeñas.

En valores estandarizados queda patente la enorme carga de la varianza condicional en el período anterior para la definición de la varianza condicional en el momento actual (el parámetro supone el 75% de las variaciones otorgadas a la generación de dicha varianza condicional).

En cuanto al modelo en media, los parámetros del modelo de medias móviles permanecen, lógicamente, casi invariados, percibiéndose un cambio importante en el valor del parámetro de la varianza condicional como explicativa del rendimiento medio ( $\gamma$ ) entre el primer y el segundo de los modelos planteados, permaneciendo relativamente constante a partir de entonces.



### Modelo VI: especificación con potencias asimétricas (AP-ARCH)

Hasta el momento, la modelización que hemos planteado se rige por la "inercia" lógica de la forma de cálculo de la varianza del error; es decir, el valor al cuadrado de los errores del modelo en media. Sin embargo, la alta correlación entre los valores desplazados del valor absoluto de la perturbación y ciertas potencias en torno a uno de esta serie justifican plenamente el empleo de un modelo tipo A-PARCH como sugería Ding (1993) a la hora de especificar la evolución de la varianza heterocedástica de la serie de rendimientos del IBEX-35.

Para la estimación se ha creado un programa ad-hoc sobre el paquete estadístico E-Views dado que esta formulación no existe entre los objetos de estimación ya prediseñados en dicho paquete. Las características de los valores empleados son los siguientes:

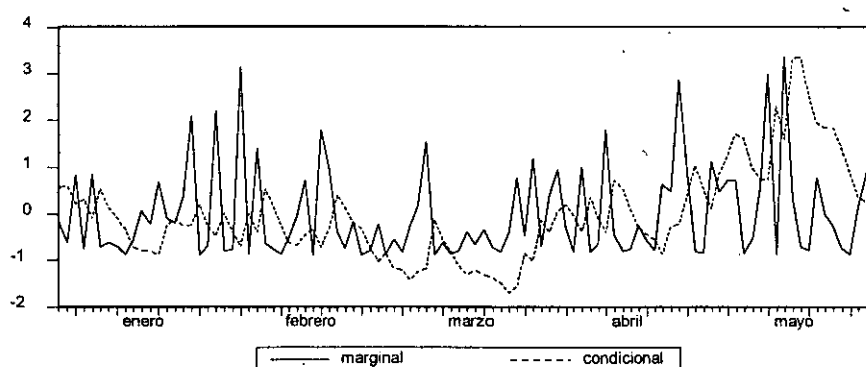
Modelo estimado	$\varepsilon_t = y_t - \mu + \theta_1 \varepsilon_{t-1}$ $h_t^\delta = \omega + \alpha_1 ( \varepsilon_{t-1}  + \gamma \varepsilon_{t-1})^\delta + \beta_1 h_{t-1}^\delta$
Período	14 de enero de 1992 - 6 de junio de 2000
Valores iniciales de los parámetros	Proceso GARCH(1,1)
Función de distribución de los residuos	Normal (gausiana)
Algoritmo de resolución	Berndt-Hall-Hall-Hausman
Estimación de la matriz de varianzas-covarianzas	Cuasi-máxima verosimilitud de Bollerslev y Woodridge (1992)

Los resultados de estimación son los siguientes:

	Parámetros		Desv. Típica	z- Estadística	Prob.
	Nombre	Valor			
<b>Modelo de la Media</b>	MU	0.000762	0.000257	2.962922	0.0015
	THETA	0.121868	0.023513	5.182925	0.0000
	OMEGA	2.96E-05	3.50E-05	0.845448	0.1989
<b>Modelo de la Varianza condicional</b>	ALPHA	0.086878	0.013696	6.343298	0.0000
	GAMMA	0.203023	0.066851	3.036934	0.0012
	BETA	0.893611	0.014662	60.94673	0.0000
<b>Potencia asimétrica</b>	DELTA	1.563314	0.250268	6.246564	0.0000
<b>Log verosimilitud</b>		6451.789	<b>Akaike info criterion</b>		-6.12041
<b>Media Log verosimilitud</b>		3.063527	<b>Schwarz criterion</b>		-6.10162
<b>Nº de Explicativas</b>		7	<b>Hannan-Quinn criter.</b>		-6.11353

En este modelo todos los coeficientes calculados son significativamente distintos de cero, salvo el término constante de la especificación de la desviación típica elevada a delta. El contraste de restricción de coeficientes de Wald no acepta la hipótesis nula de valor nulo de este término omega salvo para el 25% de las ocasiones, valor que, si bien está muy por debajo del 5% deseable, nos inclina a mantener el término por su significación económica. Para el resto de los coeficientes, la restricción de nulidad es inaceptable para valores de probabilidad superiores al 99%. (según el mismo tipo de contraste).

Gráfica de la varianza marginal y la condicional estimada en el año 2000



El valor obtenido del coeficiente añadido en el modelo A-PARCH (delta) resulta bastante más próximo al cuadrado que el resultante en la experiencia de Ding (1992); pero aún así, su valor de 1,56 y su clara significación son una llamada de atención sobre la dudosa idoneidad de realizar una especificación directamente en términos al cuadrado de la varianza. Parece que los resultados son concluyentes de lo contrario,

ya que el máximo de la función de verosimilitud se logra con un valor de este parámetro distinto a dos.

La constatación previa de la no normalidad de la serie modelizada hace necesario comprobar cuál sería el valor de la estimación empleando la forma, en principio, más adecuada; es decir, la distribución t-student para la determinación de los parámetros como maximiza

		Parámetros		Desv. Típica	z- estadística	Prob.
	Nombre	Valor				
Modelo de la media	MU	0.000794	0.000250	3.170371	0.0008	
	THETA	0.118838	0.023044	5.157032	0.0000	
Modelo de la Varianza Condicional	OMEGA	4.06E-05	6.23E-05	0.652246	0.2571	
	ALPHA	0.094436	0.016743	5.640204	0.0000	
	GAMMA	0.227925	0.082036	2.778356	0.0027	
	BETA	0.898024	0.016784	53.50594	0.0000	
Potencia asimétrica	DELTA	1.426959	0.334190	4.269907	0.0000	
	Grados libertad	11.08152	1.638363	6.763774	0.0000	
Log likelihood		6452.678	Akaike info criterion		-6.14521	
Avg. log likelihood		3.076404	Schwarz criterion		-6.12374	
Number of Coefs.		8	Hannan-Quinn criter.		-6.13735	

Se puede observar directamente que la diferencia de estimación considerando la distribución normal (aunque empleando la estimación eficiente de la matriz de varianzas covarianzas de Bollerslev y Woodridge -1992-) y la distribución t-student es prácticamente nula en los valores de los parámetros calculados aunque, el valor del logaritmo máximo verosímil obtenido es algo mayor. Donde sí se aprecia un valor notablemente más parecido al que obtenían otros autores para índices americanos o de otros países europeos es en el valor del parámetro delta o de la potencia simétrica a aplicar. Ahora estaríamos hablando de un valor de 1.43 frente al 1.58 previo.

Modelo	MSE	MAE	LMSE	LMAE	LOG MAX- VEROSIM
<b>AP-Arch</b>	<b>1,06*10<sup>-7</sup></b>	1,5410 <sup>-5</sup>	7,16	1,82	6451,792
<b>AP-Arch (t-student)</b>	<b>1,06*10<sup>-7</sup></b>	1,55*10 <sup>-5</sup>	7,16	<b>1,81</b>	<b>6452,678</b>

Haciendo uso de las funciones de pérdida ya empleadas previamente, se puede indicar que el error por empleo de una función de distribución normal en vez de una, quizá más apropiada, t de Student no parece afectar de un modo importante a estos

estadísticos, siendo muy similares los resultados obtenidos con los dos primeros modelos y habiendo una escasa ganancia en la función máximo verosímil obtenida.

### **Modelo VII: Una propuesta alternativa: modelo GJR introduciendo el óptimo de desfase de correlación tipo APARCH**

Hasta el momento, los modelos estimados nos llevan a una conclusión difícil:

Por un lado, el modelo GJR da lugar a un mayor ajuste y el seguimiento del fenómeno que nos ocupa (la volatilidad condicional) parece estar mucho más en consonancia con la volatilidad observada.

Por otro lado, el modelo sugerido por Ding (1993) tiene buena lógica en el contexto de la serie del Ibex-35, donde la evidencia de la estructura de correlaciones sobre valores con potencias asimétricas se ha puesto claramente de manifiesto. No se puede sostener una potencia cuadrada cuando la que es estadísticamente óptima está en torno a 1,5 o algo por debajo (1,42). La propia estructura de autocorrelación nos muestra que la potencia uno del valor absoluto es la que presenta el máximo considerado sólo el primer retardo.

En este contexto, se plantea la estimación de un nuevo modelo que se formularía del siguiente modo:

$$\begin{aligned}\varepsilon_t &= y_t - \mu - \theta\varepsilon_{t-1} + \gamma_1 h_{t-1}^\delta \\ h_t^\delta &= \omega + \alpha_1 |\eta_{t-1}^2| + \alpha_2 \eta_{t-1}^\delta I_{t-1} + \gamma_2 \text{Ln} \text{mibor}_t m_{t-1} + \beta_1 h_{t-1}^\delta \\ \eta_t &= (1 + \lambda_1 \text{enero}_t + \lambda_2 \text{diciembre}_t) \varepsilon_t\end{aligned}$$

En dicho modelo aunaríamos la capacidad explicativa demostrada para las variables de efecto apalancamiento, rentabilidad sin riesgos y corrección estacional (modelo GJR-ARCH) y la potencia donde la función de autocorrelación obtiene sus máximos. El modelo en media se construye como en el resto de los casos; es decir, con unas medias móviles de orden uno y la varianza condicional desplazada un período como explicativa adicional.

En este modelo se supone nuevamente una distribución gaussiana del término estocástico de error una vez ya ha sido filtrado de los efectos de estacionalidad a partir

de la regresión con las variables ficticias seleccionadas (enero y diciembre). Aún así, se estimará el modelo con la función cuasi-máximo verosímil de Wooldridge Bollerslev (1992), con lo que podríamos obviar el tema de la función de distribución.

Los resultados de la estimación, derivados del programa de e-views ad-hoc adjunto, se muestran a continuación:

	Nombre	Parámetros		z-Statistic
		Valor	Desv. Típica	
<b>Modelo de la media</b>	MU	0,000	0,001	-0,406
	THETA	0,140	0,024	5,913
	GAMMA(1)	0,053	0,010	5,422
<b>Potencia asimétrica</b>	DELTA	1,027	0,195	5,255
<b>Corrección estacional</b>	LAMBDA(2)	0,001	0,000	8,306
	LAMBDA(3)	0,002	0,001	2,060
<b>Modelo de la varianza condicional</b>	OMEGA	0,001	0,001	0,828
	ALPHA(1)	0,05	0,009	4,049
	ALPHA(2)	0,07	0,019	2,985
	BETA	0,829	0,016	50,390
	GAMMA(2)	0,002	0,000	7,972
Log likelihood		6,555	<b>Akaike info criterion</b>	-6,215
Avg, log likelihood		3,113	<b>Schwarz criterion</b>	-6,285
Number of Coefs,		11	<b>Hannan-Quinn criter,</b>	-6,204

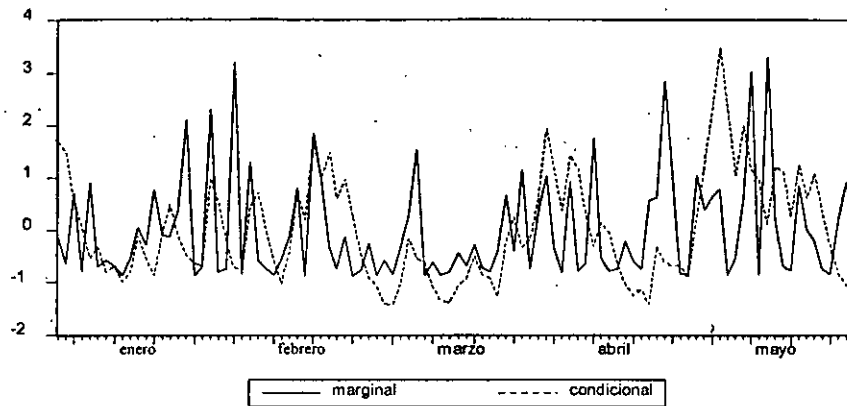
Nótese que en el modelo propuesto el valor del logaritmo de máxima verosimilitud es el mayor de cuantos se hayan presentado (en esta ocasión, la estimación se realizó sobre el valor del logaritmo de la rentabilidad sin multiplicar por 100, por lo que las comparaciones deberán hacerse sobre los modelos que permitan estimar esta variable).

Es interesante reseñar que el parámetro delta (la potencia asimétrica que enunciaba Ding) se presenta en el valor 1,027, lo que sería casi igual que decir que el valor con mayor autocorrelación es directamente el valor absoluto del error muestral, o un valor extremadamente próximo a éste.

El modelo presenta un grado de ajuste entre volatilidad real y volatilidad condicional mucho más preciso que el de los modelos anteriores. En el gráfico se presentan las últimas cien observaciones de este índice, quedando patente cómo la volatilidad condicional replica la observada con un cierto adelanto más o menos constante; es

decir, la variable varianza condicional, traducida como volatilidad que conoce el inversor en el día que opera y que, por tanto, aplica para determinar su riesgo como una explicativa más, está replicando en cierta forma la volatilidad observable previamente pero ahora de forma exacta, no aleatoria.

Gráfica de la varianza marginal y la condicional estimada



De cara a fijar las diferencias que supone este nuevo modelo, de mejor valor del logaritmo máximo encontrado, frente al modelo IV que habíamos elegido como el más completo, siguiendo la estimación tradicional tipo ARCH, presentamos el siguiente cuadro:

## Diferencias GJR - AP - ARCH y GJR-ARCH

	GJR - APARCH		GJR - ARCH		% DIFEREN
	Parámetro	z-estadística	Parámetro	z-estadística	
MU	0,000	0,001	0,000	0,000	
THETA	0,140	0,024	0,119	0,024	
GAMMA(1)	0,053	0,010	0,729	0,365	
LAMBDA(2)	0,001	0,000	0,001	0,000	
LAMBDA(3)	0,002	0,001	0,002	0,001	
OMEGA	0,001	0,001	0,000	0,000	
ALPHA(1)	0,035	0,009	0,055	0,015	
ALPHA(2)	0,175	0,059	0,060	0,018	
BETA	0,829	0,016	0,881	0,016	
CHI	0,002	0,000	0,000	0,000	
DELTA	1,027	0,195	-	-	
Log Máxima verosimilitud	6,555	-	6,454,6	-	1,6%
Media. log Máximaver.	3,113	-	3,065	-	1,6%
Nº de parámetros	11	-	10	-	10,0%
criterio informativo Akaike	-6,215	-	-6,120	-	1,5%
criterio informativo Schwarz	-6,285	-	-6,093	-	1,5%
criter. Hannan-Quinn	-6,204	-	-6,110	-	1,5%
MSE	0,8929	-	1,066	-	-16,2%
LMSE	6,4920	-	7,091	-	-8,4%
MAE	0,0001	-	0,000	-	-9,7%
LMAE	1,7193	-	1,821	-	-5,6%

El porcentaje de diferencia se calcula sobre la resta del valor en el nuevo modelo frente al GJR

El modelo nuevo propuesto presenta las siguientes ventajas respecto a los anteriores:

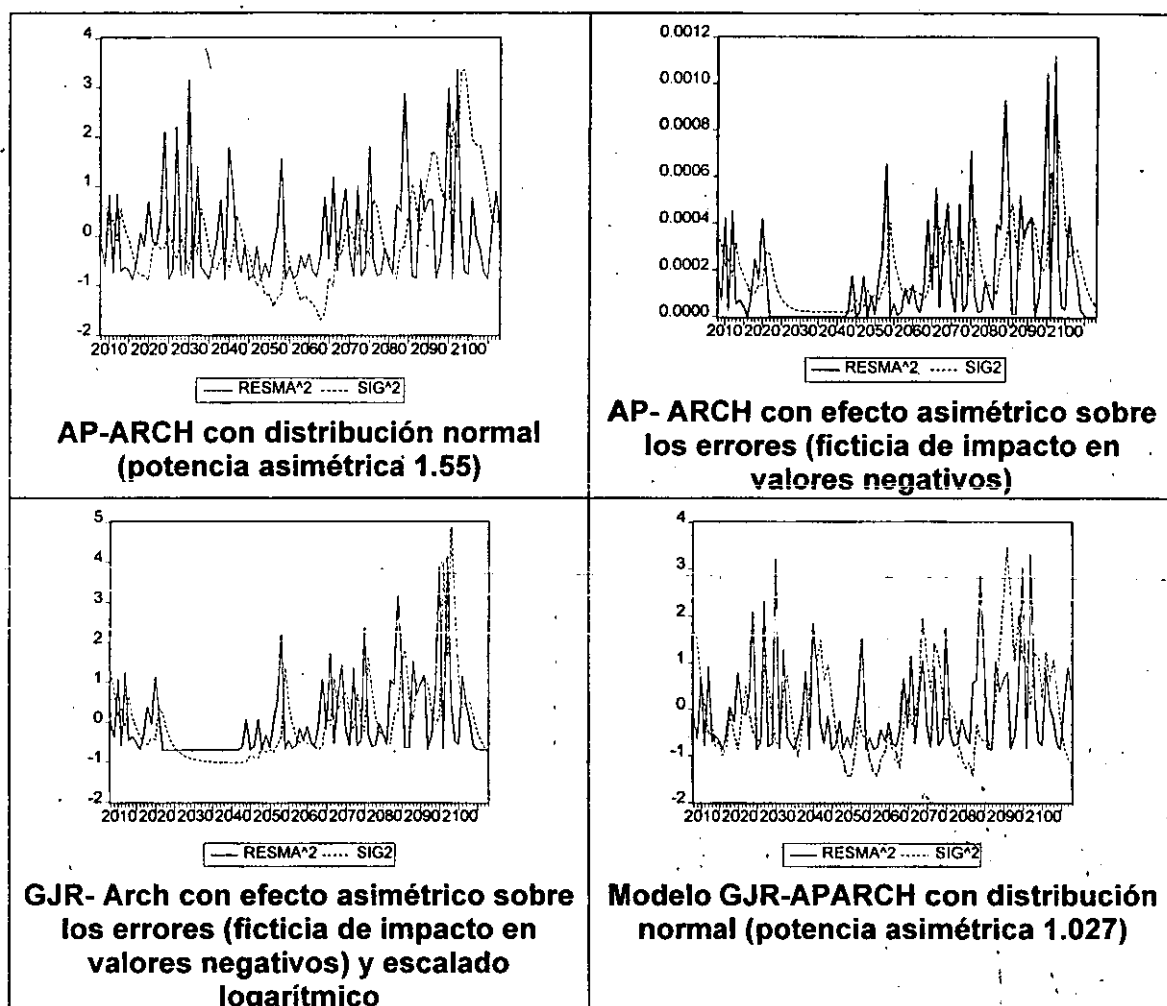
- el logaritmo de verosimilitud es un 1,6% superior al del modelo GJR-ARCH
- Los criterios de Akaike, Schwarz y Hannan- Quinn coinciden en señalar que la pérdida de grados de libertad por la inclusión de una variable más en la estimación está suficientemente justificada por la reducción en la suma cuadrática de los errores<sup>30</sup>.
- Las cuatro funciones de pérdida calculadas en todos los casos anteriores<sup>31</sup> confirman a este nuevo modelo como mejor en función del grado de ajuste entre la varianza marginal y la varianza condicional.

<sup>30</sup> En la tabla hay que tener en cuenta que son valores negativos.

<sup>31</sup> MAE: error absoluto medio, MSE: error cuadrático medio; LMAE: MAE en logaritmos y LMSE: MSE en logaritmos.

En definitiva, la propuesta que incluimos como novedad es la que representa, para el caso del Ibex-35, los mejores resultados obtenidos hasta el momento con la modelización tipo ARCH.

Gráficos de ajuste entre varianza condicional y la observada



A la luz de la poco fiable comparación gráfica, pero interesante de cara a la predicción, la diferencia en mejoría del último modelo presentado frente a los previos es evidente, corrigiéndose el período mal recogido por todos los modelos anteriores que correspondería con marzo del año 1999.

En definitiva, los resultados de esta nueva propuesta que llamábamos GJR-APARCH suponen un modelo netamente mejor en todos los aspectos a los construidos anteriormente, por lo que entendemos que éste es el que nos ofrece una buena medida de la volatilidad condicional y ésta, a su vez, una buena medida del riesgo por sí sola.

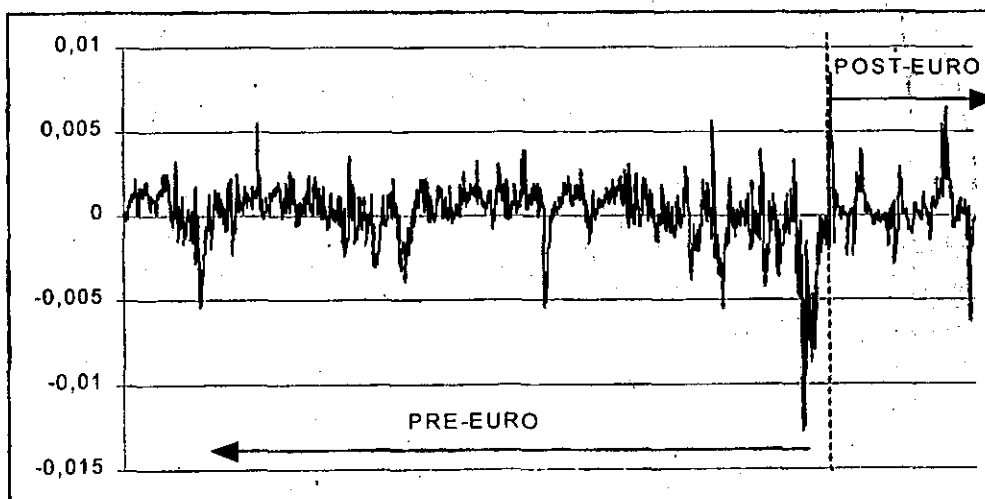


### Comparación de los modelos presentados y contraste de permanencia estructural del modelo elegido. Análisis de la volatilidad estimada

A la luz de la volatilidad calculable con el último modelo propuesto (al que hemos llamado "GJR-APARCH") tenemos una primera medida de este factor de mayor o menor premio exigible a un activo financiero en bolsa a partir del mayor o menor grado de incertidumbre al que dé lugar.

En cualquier caso, previamente a la realización de algunas observaciones sobre la evolución de la varianza condicional estimada, cabe un último planteamiento sobre la permanencia estructural del modelo planteado. Para ello, se ha comprobado si habría o no diferencias notables en los valores de los coeficientes estimados cuando se realiza una estimación de la varianza con el modelo elegido en dos submuestras diferentes: antes de la creación oficial del Euro (1 de enero de 1999) y después de dicha creación. Salvando los sesgos que la enorme deferencia en el número de observaciones disponibles pudieran ocasionar, un primer contraste gráfico nos deja entrever que, efectivamente, las diferencias de estimación son notables, pero tan sólo en puntos inmediatamente anteriores o posteriores a la puesta en vigor de la moneda única europea.

#### Diferencias en la estimación para el total de la muestra y por submuestras



En el gráfico, la línea discontinua parte ambos periodos, siendo relevante que las mayores discrepancias que muestran las estimaciones por subperiodos frente a la

estimación con la muestra completa se producen en los días aledaños a la instauración del Euro como moneda única, señalando quizá los momentos en los que los mercados pudieran estar anticipando y corrigiendo sus expectativas sobre la creación de un área con una sola moneda. Este hecho, en la estimación con los parámetros de la muestra completa, queda mucho más diluido, ofreciéndose un valor estimado de la volatilidad mucho más pequeño.

Como primera medida estadística sobre la conveniencia de considerar un cambio de estructura en la serie, cabe plantearse la suma de los errores al cuadrado que, aunque de un modo muy deficiente por no contar con una distribución para la contrastación paramétrica, nos darían una idea aproximada. En este caso, los errores cuadrados medios son 7,82, 6,84 y 6,49 para los períodos pre-euro, post-euro y total respectivamente, con lo que parece que, en niveles bastante similares, el más conveniente seguiría siendo el modelo total, no por submuestras.

Como una segunda alternativa de medición aproximada de los cambios en la volatilidad estimada, se ha calculado la media de la volatilidad en el período de existencia del Euro y la del período anterior, siendo significativo que este segundo período (post-euro) mostraba un valor medio superior a un 13% a la volatilidad media total, mientras que la época pre-euro mostraría una media en volatilidad muy homogénea como la que resulta del cálculo con la muestra completa. A pesar de lo dicho, el desproporcionado número de observaciones entre un período y otro y la contribución de cada uno de ellos a la media total hacen necesarias algunas medidas alternativas.

Se calculó la media de volatilidad estimada año y medio antes de la creación del euro y año y medio después (fin de la muestra). El resultado es absolutamente contrario al anterior y, en este caso, la divergencia del período previo respecto a la volatilidad media de todo el período resulta ser de un 23%; frente al 13,5% que habíamos calculado para la época post-euro. Según este cálculo, la volatilidad se habría suavizado intensamente después de la creación de la moneda única.

Para descartar el que la creación de la moneda única pudiera estar delimitando un cambio de estructura de los llamados de "punto conocido", planteamos un análisis como el sugerido por Shiang (1995), de forma que calculamos el estadístico experimental tipo multiplicador de Lagrange de distribución teórica conocida:

$$LM(\delta) = T \left[ \frac{\partial L(\delta)}{\partial \theta'} \right] B_0^{-1} \left[ \frac{\partial L(\delta)}{\partial \theta'} \right] \Rightarrow \chi_{p+q+1}^2$$

donde

$$B_0^{-1} = \frac{1}{4} \begin{bmatrix} \delta v_0 & 0 \\ 0 & (1 - \delta v_0) \end{bmatrix}$$

y

$$v_0 = h_{0t} \left( \frac{\partial h_{0t}}{\partial \theta_1} \right) \left( \frac{\varepsilon_{t-1}^2}{h_{0t}} - 1 \right), \quad h_{0t} = \theta_0' Z_t$$

El valor obtenido de dicho contraste confirma que no existiría un cambio de estructura en el punto señalado con una significatividad de más del 98%.

Dicho esto, no analizaremos la posibilidad de generar un modelo tipo Switching ARCH, dada su complejidad iterativa y que parece quedar demostrado que no se producen diferentes estados, por lo menos de una forma continuada desde su potencial establecimiento.

Ante esta situación, la volatilidad estimada con el llamado modelo GJR-APARCH, que damos por útil, se convierte en el valor de referencia en dos sentidos:

- Nos permite relacionar directamente la varianza condicional estimada con el valor de la rentabilidad exigida al Ibex-35 a partir del parámetro gamma. Sin embargo, la compleja estimación sobre la diferencia logarítmica hacen que el cálculo no sea inmediato y que esté fuertemente sujeto a ligerísimos cambios.
- Siguiendo los enunciados de la teoría de la CAPM y APT combinadas, por cuanto se ha estimado el rendimiento en función de la volatilidad y de una serie de factores adicionales (corrección estacional y activo sin riesgo), obtenemos el valor de "medición de la incertidumbre" que se define en estas teorías y en otras como la volatilidad condicional estimada. La ventaja de esta metodología de cálculo, además de su mayor exactitud puntual, es que obtenemos una serie con variación en cada momento del tiempo que resulta ser igual al ajuste que los agentes económicos realizan.

Aún cuando es evidente la conexión entre rentabilidad y volatilidad a partir del modelo de regresión, siguiendo el hábito en la utilización de "indicadores de alerta"

entre los operadores de bolsa más de tipo cualitativo , se propone la determinación de uno nuevo con la distribución de percentiles de la volatilidad, por ejemplo, para cada año, acotando así períodos en función de la volatilidad observada y de predicción y dando lugar a señales cualitativas de alerta ante determinados valores.

En el caso de la volatilidad estimada para el Ibex-35 los resultados de esta distribución por decilas sería la siguiente en el año 2000<sup>32</sup>:

Intervalos de la volatilidad estimada en el año 2000

	Percentil Real	Percentil Uniforme	Diferencia
10%	1,12	0,91	0,21
20%	1,18	1,05	0,12
30%	1,24	1,19	0,05
40%	1,31	1,34	-0,03
50%	1,39	1,48	-0,09
60%	1,48	1,62	-0,14
70%	1,57	1,76	-0,19
80%	1,69	1,91	-0,22
90%	1,77	2,05	-0,28
100%	2,19	2,19	0,00

Se emplea el valor del percentil real entendiendo que las diferencias en la distribución de frecuencias observables a partir de la distribución en diez partes iguales del recorrido (percentil uniforme) son suficientemente grandes como para desaconsejar este segundo método.

En el anexo I se recogen los resultados de la volatilidad estimada "valorados" de uno a diez a partir del percentil que le corresponda a la información para un año completo en cada momento (el percentil de las últimas 250 observaciones, ya que son días hábiles). Se han impreso tan solo los valores recogidos desde la creación del Euro, para evitar una tabla de los más de 2000 valores muestrales. Ello nos permite dar una medida más intuitiva del fenómeno, ya que lo referenciamos a una escala decimal.

<sup>32</sup> La volatilidad estimada está aquí multiplicada por 100.

Aplicando los límites comunes en bolsa para un índice tipo RSI<sup>33</sup>, se ha clasificado cada valor en función de si está en la decila por debajo del 20% o por encima del 80%, entendiendo respectivamente si estamos ante una situación de altísimo riesgo o de riesgo muy bajo.

El primer resultado interesante en si mismo sobre la volatilidad estimada, por cuanto viene a confirmar diversos planteamientos teóricos con significatividad estadística en los últimos diez años, es que la varianza condicional estimada:

- está presente en el proceso de formación de la rentabilidad de un activo bursátil como el Ibex - 35 y es un punto a tener en cuenta para medir la prima de riesgo.
- Dicha prima de riesgo es directamente proporcional a la volatilidad; es decir cuando aumenta la varianza condicional, aumenta la rentabilidad exigida al Ibex-35.
- Sigue habiendo un retardo en la aplicación de resultados complejos a las exigencias del día, y "la varianza condicional que sirve para calcular la rentabilidad de hoy es la que se calculaba ayer".

En el proceso de estimación, el empleo de la matriz de varianzas covarianzas formulada por Woodridge y Bollerslev (1992) nos permite salvar el problema de desconocimiento de la función de densidad real de la variable a modelizar y nos ofrece unos resultados consistentes. De hecho, la estimación a partir de funciones de densidad más parecidas a la representación muestral del error como pudieran serlo la t-student o la GED propuesta por Nelson (1992) se han demostrado inútiles para mejorar los resultados y han ofrecido un valor de los parámetros estimados muy próximo a lo que obteníamos con el método propuesto.

La varianza condicional estimada y la varianza marginal (calculada directamente como cuadrado de los residuos esta última) muestran un cierto comportamiento tipo réplica de la marginal por la condicional con un retardo. Es decir, parece que la variable que

<sup>33</sup> El RSI (*Relative Straiigth Index*) es un popular mecanismo de seguimiento de los valores bursátiles que permite al agente valorar cuando un activo está siendo sobrevendido o sobrecomprado, llevando al tenedor a comprarlo (por estar en su nivel mínimo de precio) o venderlo (por estar en su nivel máximo) respectivamente. Este índice basa sus

se muestra significativa en la estimación del rendimiento del índice bursátil es un conocimiento aproximado de la varianza calculable en el período anterior más un cierto componente añadido por los errores que en este conocimiento se produjeron. Aún así, es obvio que el seguimiento de la varianza marginal por la condicional es deficiente y se hallan, en el gráfico, multitud de puntos en los que el signo de los puntos críticos es exactamente el contrario.

Los resultados están conformes con el hecho estilizado que ya ha sido numerosas veces referido de períodos de calma y períodos de alta volatilidad distribuidos en el tiempo de forma agrupada.

Contaríamos entonces con una herramienta que, con un complejo trasfondo, se convierte en algo fácil de utilizar como señal de alerta y o prevención de riesgo.

---

recomendaciones en función de cuando se sobrepasa un 75% de los precios históricos del último año.

## **9. CONSIDERACIONES FINALES**

En el trabajo que presentamos se ha realizado un exhaustivo recorrido y análisis a través de los dieciocho años de existencia de los modelos ARCH, exponiendo con rigurosidad las sucesivas aportaciones metodológicas a las que su intensiva utilización han ido dando lugar durante todos estos años.

En el artículo seminal de Engle de 1982, se planteaba la determinación de un proceso generador de datos que se muestra incompleto cuando la estacionariedad se verifica tan sólo en términos marginales, obviando una estructura significativa de correlaciones en los segundos momentos de la serie (la varianza), que necesariamente deben ser tenidos en cuenta para la determinación final del proceso de formación de una serie temporal. Ante esta situación, se construye un sistema metodológicamente más completo en la especificación de una serie autoalimentada por su pasado. Las aportaciones de los años inmediatamente posteriores irán por el camino de completar y mejorar esta definición estadística (GARCH, ARCH en media, IGARCH, etc.)

A partir de entonces, llama particularmente la atención cómo la realidad y la lógica más básicas de un mercado financiero encuentran su equivalencia estadística con los modelos ARCH casi como si de un ejercicio de traducción se tratara. A partir de un esquema tan específico como el que planteaba Merton (1980) para la previsión del rendimiento de un activo financiero y la valoración de las opciones, la traslación de la incertidumbre (la expectativa) sobre el riesgo no medible a la esperanza de la volatilidad medida como varianza es un paso evidente; como también lo es que en la formación de expectativas tiene un peso fundamental la información recibida del pasado y el error cometido en dicho pasado al preverla.

El intensivo empleo de los modelos tipo ARCH durante los años siguientes genera un gran número de aportaciones accesorias que, posteriormente, se derivarían estadísticamente (teóricamente) y no al contrario. Desde 1986, la evidencia empírica modelizada como una estructura ARCH, aunque incorporando nuevos términos en la especificación, va dando lugar a una incesante cascada de nuevas incorporaciones metodológicas que en esta tesis hemos venido a llamar variantes sobre los modelos ARCH.

A partir de los años 90, la corriente más *teorista*, de la que son buenos representantes Engle, Bollerslev, Nelson, Gouriéroux o Hamilton, convive con un importante elenco de *aplicacionistas*, sobre todo a partir de las aportaciones realizadas



desde diversas Reservas Federales estadounidenses. Glosten, Jagannathan, Runkle, Susmel, Schwert, Ding e, incluso, Granger son buenos representantes de esta segunda tendencia que podríamos identificar con un proceso "de lo aplicado a lo teórico". Estaríamos, en este segundo caso, en un desarrollo "inductivo" frente al más común proceso "deductivo" que han experimentado otros planteamientos econométricos.

En esta tesis, se ha pretendido, partiendo de un análisis estadístico-econométrico riguroso, plantear un desarrollo modelizador muy similar al que la propia historia de los modelos ARCH nos ha venido ofreciendo. Intercalando convenientemente teoría y práctica, se han puesto de manifiesto las debilidades y fortalezas de los diversos métodos presentados y su adaptabilidad meramente práctica a la variable que hemos escogido como objetivo, el IBEX-35.

En el marco descrito, sería conveniente realizar algunas consideraciones específicas que se agruparían en torno a cuatro grandes líneas:

- I. Adecuación de las distintas especificaciones alternativas a las realidades empíricas observables en los mercados financieros.
- II. Utilidad final de los modelos tipo ARCH para la determinación de los mercados financieros y nueva aproximación tipo ARCH planteada.
- III. Consideraciones sobre la metodología econométrica.
- IV. Otras posibilidades en el entorno de los modelos de varianza condicional y futuras líneas de investigación.

**I.- Adecuación de las distintas especificaciones alternativas a las realidades empíricas observables en los mercados financieros.**

La ventaja de los modelos tipo ARCH para recoger elementos económicos recurrentes en las series financieras (*stylized facts*) es evidente. Factores hartamente estudiados encuentran claro reflejo en este tipo de modelos:

- Por un lado, la importancia de la volatilidad en la medición del riesgo y las implicaciones de este a la hora de determinar el diferencial de rentabilidad que se le exige a un activo no seguro sobre el activo público libre de riesgo.
- En segundo lugar, los factores de "contagio financiero" (períodos volátiles vienen precedidos de otros del mismo modo y períodos tranquilos de períodos del mismo tipo) quedan perfectamente descritos con estos modelos no lineales.
- Por último, el carácter asimétrico de la respuesta ante sobrevaloraciones o infravaloraciones de los rendimientos esperados también queda plenamente modelizado.

A juicio de Rich y otros (1991), los modelos ARCH son tan profusamente utilizados en finanzas fundamentalmente por dos razones:

- En primer lugar, se centran en la estimación de la evolución temporal de la varianza condicional, directamente entroncada con la parte no predecible de las series temporales así como mejor medida del riesgo y de la incertidumbre que otras "proxis" con base en otras medidas de la volatilidad.
- En segundo lugar, las propiedades estadísticas de los modelos ARCH parecen ser una descripción exhaustiva y adecuadamente parca (*parsimonious*) del comportamiento de muchas series financieras.

De hecho, la preocupación de la modelización tipo ARCH por contemplar algunas situaciones evidentes en el funcionamiento de las variables financieras ha encontrado un nicho metodológico de extraordinario desarrollo. Parcelas empíricas como la necesaria corrección estacional del residuo ante las evidentes diferencias de los meses y días de la semana por su volatilidad observada eran un punto que necesariamente debía incluirse y sistematizarse en la generación de la varianza condicional. El agente económico, al formar sus expectativas y generar una determinada volatilidad, tiene en cuenta el pasado ya conocido (principio en el que se fundamenta el desarrollo teórico planteado) y, además, es consciente de diferentes situaciones que pueden afectar al rendimiento en función del mes del año en que nos encontremos.

Otro carácter de crucial importancia en los modelos que planteamos, y más para los mercados en los que se utilizan, es, como decíamos, el efecto asimétrico llevado a la generación de la volatilidad. Es un hecho constatado que ante una infravaloración de un rendimiento, el agente es cauteloso y se va convenciendo despacio y progresivamente de este hecho y su mantenimiento en el tiempo. Sin embargo, ante una sobrevaloración del rendimiento, la reacción correctora es casi inmediata y de forma drástica. Dicho esto, la discriminación en los resultados de conocer una información del pasado sabiendo también su signo es una necesidad clara en la modelización de los rendimientos de los activos financieros a la que los modelos planteados responden de un modo adecuado. El ajuste a la realidad con la inclusión de una variable que discrimina según el signo del error cometido en el período precedente se ve enormemente mejorado.

Este carácter asimétrico en la volatilidad se hace patente desde el primer momento en las diversas modelizaciones del IBEX-35 realizadas. El hecho contrastado empíricamente de que una sobrevaloración del rendimiento en el período precedente da lugar a una reacción más fuerte que una infravaloración se ha venido poniendo de manifiesto desde la primera especificación propuesta y la significatividad estadística de incluir una variable ficticia discriminante de esta situación viene confirmada por los contrastes de sesgo de signo realizados.

Un resultado interesante en cuanto al carácter asimétrico es que el impacto en la volatilidad actual de una infravaloración en el período precedente es de más del doble. A la luz de los parámetros estimados, los momentos de sobrevaloración incorporan 0,05 puntos por el valor del error; mientras que los de infravaloración, además de estos, incorporan 0,07 puntos. Es decir, un total de 0.12 (infravaloración) frente a 0,05 (sobrevaloración).

## **II.- Utilidad final de los modelos tipo ARCH para la determinación de los mercados financieros y nueva aproximación tipo ARCH planteada**

Con estos condicionantes de partida parece, por tanto, adecuado recurrir a una estimación de la varianza condicional para la buena identificación de unos procesos, como los reflejados en los mercados financieros, cuyas estructuras de autocorrelaciones en los valores absolutos, o potencias, presenta valores significativos.

Por supuesto, para la estimación paramétrica el conocimiento exacto de la verdadera varianza de la distribución que se intenta maximizar aumentará la consistencia de los coeficientes calculados.

Para evaluar la potencia de este tipo de especificaciones, se ha seleccionado una serie estadística de carácter claramente financiero y, por tanto, cuyo proceso generador se deriva de la apreciación del riesgo, elemento de difícil medición cuya aproximación cuantitativa supone la mayor aspiración de los agentes económicos dedicados a las finanzas desde que Bachellor, en 1900, enunciara su principio de valoración de las opciones, recogido en su tesis doctoral (Université de la Sorbonne, Paris 1900).

En este marco, se plantea una modelización relativamente novedosa en el largo recorrido de la econometría empírica, cuyo estudio consideramos de una importancia extraordinaria por dos grandes razones. En primer lugar, porque la idea seminal de Engle de 1982 permitía aunar casi de forma intuitiva términos estadísticos con evidencias financieras lógicas de cara a dotar de un mecanismo apropiado y estadísticamente correcto el funcionamiento de las previsiones basadas en los riesgos inherentes a la volatilidad de un activo. En segundo lugar, porque los mercados financieros emplean este tipo de modelos con habitualidad, por lo que conocerlos con detenimiento es una forma de prever los movimientos en este tipo de mercados, al tiempo que adaptarlos mejor a la realidad puede contribuir a dotar de mayor eficiencia nuestras decisiones en esta materia.

Para la aplicación empírica desarrollada, se ha prestado especial atención a la aproximación hecha por Glosten, L. Jagannathan, R. Runkle, D. (1993), que incorporan al modelo habitual el carácter de los factores estacionales cuya evidencia empírica es innegable, así como los avances en la técnica logrados hasta el momento. En el análisis econométrico, se pone de manifiesto la importancia del mes en el que nos encontramos para determinar un comportamiento más o menos volátil. Con el análisis descriptivo y el análisis estadístico, se encontró claramente significativo tener en cuenta los meses de enero y diciembre por su hecho diferencial en como causa de la volatilidad.

Los resultados del modelo tipo GJR-ARCH aplicado nos han permitido observar un buen seguimiento de las puntas y valles de la volatilidad observada y el modelo AP-

ARCH la correlación potencial asimétrica. Con el fin de unir las buenas características de ambos modelos, se ha procedido a la estimación de un modelo GJR-APARCH que da lugar a resultados mejores que con cualquiera de los dos anteriores de modo separado.

A partir de los núcleos teóricos expuestos en los modelos AP-ARCH y GRJ-ARCH, se presentó una nueva aproximación práctica que depara, a la luz de la aplicación presentada, unos resultados aceptables. Dicha aportación se podría resumir en los siguientes puntos:

- El planteamiento del modelo de media está basado en los supuestos de la CAPM, explicando las variaciones de un activo a partir de su volatilidad. Por supuesto, la especificación de la media se completó con el proceso de medias móviles subyacente en la estructura de la serie de los rendimientos del Ibex-35.
- De los modelos tipo GJR-ARCH, la propuesta se hace eco de la diversidad de factores explicativos tomados, léase la corrección estacional, los factores asimétricos, etc.
- De los modelos APARCH se recoge la idea de modelizar un valor diferente al del cuadrado ante la evidencia empírica de correlaciones superiores en potencias distintas de dos.

El modelo propuesto es óptimo tomando una potencia de valor 1,027 sobre el valor absoluto del error observado en la modelización de la media. Esta potencia es casi igual a realizar el estudio solamente sobre el valor absoluto y, sin duda, es más pequeña que la que resultaba en otras investigaciones sobre índices bursátiles de otros países. Esta particularidad en el valor de la potencia elegida para maximizar la función máximo verosímil ya se hacía patente desde el estudio de la estructura de correlaciones en la serie original del rendimiento del Ibex-35, donde el valor de correlación más alto para el primer retardo de la serie se daba precisamente en el valor absoluto de la serie, sin elevar a ninguna potencia. Como ya se comentaba, es cierto que este hecho cambiaba cuando se tomaban retardos superiores. Ello explicaría el resultado tan próximo a uno, pero no exactamente uno, que se produce en el modelo propuesto finalmente.

En otro orden de cosas, el fenómeno de persistencia en la volatilidad se sigue produciendo en este modelo y, probada una especificación tipo ARCH de componentes, los resultados no aconsejan esta formulación alternativa. Esta situación se ha venido dando en casi todas las aplicaciones de los modelos ARCH, tal y como se refiere en algunas recopilaciones ya citadas.

En cualquier caso, es importante señalar que estamos ante una alternativa entre otras formas de estimar la volatilidad, entre las que destacan:

- Estimación mensual de la volatilidad media como media de los errores muestrales al cuadrado. Este método presenta las desventajas de parámetros inconsistentes, ya que el modelo entre la media y la varianza no es bloque diagonal<sup>34</sup>.
- Estimación ARMA de la desviación típica a partir de los valores absolutos del error muestral.
- Estimación a partir de la volatilidad implícita en las opciones.
- Precios máximos y mínimos interperíodo.
- Calculando la dispersión existente que se produce al realizar diferentes encuestas o muestreos, o consultando a diversos expertos.
- En mercados con activos relativamente homogéneos, algunos autores proponen la el cálculo de la dispersión entre los diferentes precios estandarizados.

Quizá, además de la consistencia, la mayor ventaja que muestran los modelos ARCH frente a estas alternativas se encuentra en el mejor aprovechamiento de la información y de su carácter de continuidad, que en los métodos expuestos se centran en medias u observaciones puntuales, salvo en el caso del modelo ARMA. Sobre este, parece que la elección de la desviación sobre el valor absoluto de una estimación que no lo tiene en cuenta de forma simultánea incurre en los mismos problemas asintóticos que la estimación en dos etapas criticada para la primera alternativa planteada: los resultados no son consistentes.

### III.- Consideraciones sobre la metodología econométrica.

Una vez se ha puesto de manifiesto la adecuación de la modelización tipo ARCH y de su conexión con las variables objetivo sobre las que hemos diseñado esta aplicación

---

<sup>34</sup> Ver Pagan (1984)

práctica, es interesante reflexionar inicialmente sobre la forma funcional elegida. Con ello, hacemos referencia a un planteamiento tan viejo como los propios modelos ARCH, por cuanto el mismo Engle (1982) ya propone en su artículo pionero modelizaciones del logaritmo del valor absoluto del residuo o del mismo valor absoluto de la serie. Sin embargo, parece que este hecho deja de ser motivo de investigación durante un largo período de tiempo, dedicándose las aplicaciones a estudiar únicamente modelos basados en la forma de la varianza: es decir el cuadrado.

Nelson (1991) despierta de nuevo la discusión y el interés por el tema cuando nos pone en guardia sobre la propagación de un sistema de estimación que, con frecuencia, se utiliza sin las debidas precauciones. La estimación "en serie" de modelos ARCH en los mercados financieros degenera su contrastación hasta el punto en el que, su utilización a cortísimo plazo, no presta especial interés a unos valores de los parámetros que originan valores de la varianza negativos cuando se amplía el horizonte de predicción.

En los años noventa, Ding y Granger trabajan nuevamente en el campo de la determinación de cuál es la forma funcional sobre la que se debe especificar el modelo o, más bien, se plantean por qué necesariamente se debe especificar la varianza sobre valores al cuadrado en el período de estimación cuando multitud de experimentos muestran estructuras de correlación del valor absoluto de los residuos con potencias diferentes de dos mayores que los que se obtienen para éstas. La lógica del planteamiento lleva a generar un tipo de modelos de potencias asimétricas que servirá de base a la nueva aproximación que se propone al final del análisis técnico de la serie del Ibex-35 y que pretende aunar y todos los factores positivos que incorpora la teoría estudiada.

Una vez constatada la existencia de autocorrelación en valores transformados del residuo estacionario de una regresión (valores absolutos, potencias de este, cuadrados, etc.), y desde el punto de vista estadístico, sería muy positivo conocer la forma en la que se distribuye esa serie del residuo transformada. Aunque la forma más fácilmente reproducible es la de la normal, varias situaciones nos llevan a pensar que la función de distribución de la serie de los momentos de órdenes superiores debe ser diferente:

- El primer indicio proviene de la contemplación del histograma de frecuencias, que resulta difícilmente aproximable a una normal. Los coeficientes de curtosis

y de asimetría también son indicios razonables de la no normalidad de la serie, circunstancias ambas que se ponen de manifiesto en el análisis técnico realizado sobre la serie del IBEX-35.

- El segundo elemento que nos hace cuestionarnos la función de densidad proviene de la consideración normal de los residuos originados en la regresión inicial. Una transformación potencial de dichos residuos por definición no puede seguir compartiendo una función de densidad de este tipo por ser una transformación no lineal.

Por supuesto, para la estimación paramétrica el conocimiento exacto de la verdadera función de distribución que se intenta maximizar aumentará la consistencia de los coeficientes calculados, y, tal como se ha comentado en varias ocasiones, la forma funcional de la densidad condicional y marginal de las variables financieras es claramente no normal: es ligeramente asimétrica y mucho más apuntada (leptocúrtica) que lo que correspondería a una distribución gaussiana.

Desde la aparición del artículo de Engle (1982) este hecho ha sido motivo de fuerte controversia y la búsqueda de la forma más correcta de estimar los parámetros de un modelo tipo ARCH ha sido objeto de diversas investigaciones, siendo de destacar las siguientes:

- Estimadores por el Método Generalizado de los Momentos, propuesto por Rich y otros (1991). Estos estimadores manifiestan algunas ventajas evidentes (no requieren un conocimiento exacto de la función de distribución de las variables, su sistema resolutivo es más fácilmente tratable que el tedioso recurso iterativo de la máxima verosimilitud, Hansen y Newey proveen de un contraste de especificación intrínseco a la estimación y extremadamente útil, etc.). Desgraciadamente, el sistema cuenta con la desventaja de ser ineficiente frente al de máxima verosimilitud cuando la verdadera función de distribución es efectivamente una normal.

- La estimación de una función GED y las disquisiciones sobre la conveniencia de la t-student propuestas por Nelson (1991) para la determinación de los parámetros son conclusión lógica de la estructura muestral del error. Sin embargo, en trabajos como el de Engle y González-Rivera (1991) se demuestra que la ineficiencia relativa de emplear la función normal cuando la verdadera función de distribución es la t-student es muy pequeña.



- Con frecuencia se han propuesto estimaciones no paramétricas de los modelos tipo ARCH empleando *Kernels* o cadenas de Fourier. De ello se hacen eco Pagan y Schwert (1991). El primero (empleando *kernels*) plantea una media ponderada en la elección del estimador de la varianza condicional donde, la ponderación, se hace depender de la historia existente sobre el fenómeno; es decir, de sus momentos condicionales en el extremo de equiponderación para todas las observaciones pasadas. La propuesta sobre la matriz de ponderaciones se realiza a partir del Kernel Gaussiano de orden uno multiplicado por la matriz de diferencias de orden determinado de las observaciones hasta el período anterior al considerado en cada momento.
- En el caso de las cadenas de Fourier, se define el llamado estimador no paramétrico flexible de Fourier definido por Gallant (1981), donde la varianza condicional se representa por polinomios de bajo orden y términos trigonométricos contruidos a partir de la información disponible en el período  $t$  sobre el pasado de la serie.
- Los estimadores propuestos hasta el momento basan sus resultados en la consistencia como principal objetivo, aunque sus autores suelen hacer hincapié en la pérdida de eficiencia que en ellos se produce. Centrando la importancia en dicha eficiencia de los estimadores obtenidos, destaca la aportación de Drost y Klaassen (1997). Estos autores basan su propuesta en el teorema LAN para los modelos de series temporales, logrando un método de estimación que no necesita imponer ninguna condición en los momentos de la serie.
- La evidente posibilidad de escribir los modelos como un modelo de parámetros cambiantes es puesta de manifiesto por Juan del Hoyo (1992) como posible estimación del modelo ARCH a partir de la representación del proceso en el espacio de los estados, poniendo como ejemplo una aplicación del filtro de Kalman del modelo de coeficientes variables implícito en el modelo ARCH(1).
- Bollerslev y Wooldridge (1992) proponen una estimación cuasi-máximo verosímil, demostrando que para una determinada clase de modelos dinámicos parametrizados por los primeros y segundos momentos, el estimador cuasimáximo verosímil es asintóticamente normal. Los autores realizan una serie de experimentos de Montecarlo sobre los que concluyen la ventaja de este tipo de estimadores (medida por el sesgo cometido) frente a los estimadores máximo verosímiles. La conclusión fundamental es

que los estimadores cuasi-máximo verosímiles aplicados a los modelos ARCH son consistentes aun cuando la verdadera función de distribución no sea la normal.

Como resumen de todos los métodos de estimación propuestos (y quizá de ello se derive la utilidad de los modelos ARCH) su capacidad para explicar el componente de incertidumbre que valoran los agentes financieros cuando determinan la exigencia de rentabilidad de un activo concreto es una variable que se puede construir de forma claramente significativa a partir de un supuesto de normalidad en la distribución empleada para la estimación de los parámetros y, siempre y cuando, se cuente con una muestra suficientemente amplia y se emplee el método de estimadores cuasi-verosímiles.

Como ha quedado de manifiesto en multitud de aplicaciones tipo ARCH realizadas desde 1982<sup>35</sup>, la estimación a partir de valores de la t-student o de la de Distribución Generalizada de los Errores propuestos por Nelson (1991) no mejoran significativamente los resultados obtenidos en la modelización de series de tipo financiero, dando lugar a una mayor complejidad en la estimación (cuando menos, exigen la estimación de un coeficiente más: los grados de libertad del modelo).

En el caso concreto del modelo seleccionado en esta tesis como mejor estimador de la varianza condicional, o del nivel de la incertidumbre de los mercados financieros que ésta representa, los resultados de emplear una distribución distinta a la normal se han demostrado inútiles en casi todos los casos, sin apreciarse ningún tipo de mejoría a juzgar por los contrastes de bondad practicados. Especial mención exige el hecho de que la modelización en logaritmos propuesta en el modelo E-GARCH tampoco ha supuesto una mejora significativa. Sobre este hecho, quizá la circunstancia de modelizar la diferencia de los logaritmos de la serie original haya dado lugar a una corrección previa que ha invalidado la presunta mejoría que desde el campo teórico se le adjudica a los modelos exponenciales.

Finalmente, y desde un punto de vista metodológico, conviene resaltar:

- por una parte, que los sistemas de estimación de los modelos ARCH, a la fuerza iterativos a partir de un algoritmo de *Scoring* (el más habitual) o, de un modo algo más correcto, de un algoritmo tipo Berndt-Hall-Hall-Hausman, son

---

<sup>35</sup> Ver Bollerslev (1994)

extremadamente sensibles a los valores de inicialización en dichos mecanismos repetitivos.

- Y por otra, y como un defecto adicional de este tipo de modelos, es necesario recalcar como las investigaciones se han centrado en la búsqueda de un estimador asintóticamente consistente, mientras que a la eficiencia del estimador no se le ha prestado especial atención.

#### **IV.- Otras posibilidades en el entorno de los modelos de varianza condicional y líneas de investigación futuras**

A pesar de que la primera aplicación de los modelos tipo ARCH se hacía sobre la serie de la inflación en el Reino Unido, son los empleos en el campo del seguimiento de las series financieras los que han marcado la pauta en la evolución de este tipo de especificaciones. Sin embargo, no hay que olvidar otras posibilidades que nos brinda el cuerpo metodológico expuesto.

En primer lugar, la estimación consistente de la varianza es un logro en este tipo de modelos y, de cara a investigaciones futuras, conjugar este hallazgo con una consecución de métodos tendentes a lograr la eficiencia es un paso adelante en el terreno de la precisión de cuantos modelos se basen en funciones de distribución en las que esté implícita una valoración de la varianza.

En diversos artículos aparecidos recientemente, la técnica aplicada en esta tesis sirve de apoyo y complemento en otros métodos de estimación, tales como los modelos VAR o los modelos de integración fraccionada tipo FI-GARCH. En el campo de los modelos multivariantes, los modelos ARCH se muestran especialmente útiles para la estimación de la matriz de covarianzas de los estimadores, tal y como se podía deducir del desarrollo expuesto sobre los modelos APT y su conexión con los modelos ARCH.

En el terreno de la modelización aplicada a las series macroeconómicas, la valoración de la incertidumbre es un factor explicativo de diversas variables, tales como la inversión, la formación de los precios, la determinación de los tipos de cambio, la evolución del empleo, la valoración de los salarios, los indicadores de previsión de

cartera, de pedidos, etc. En todas estas variables, un conocimiento mejor del mecanismo de formación de las expectativas seguramente resultará útil para definir las.

Por supuesto, y como ya ha quedado puesto de manifiesto en varias ocasiones, los modelos ARCH presentan ventajas especialmente significativas en la modelización de series de alta frecuencia, tanto por sus buenas propiedades asintóticas como por su capacidad razonable para predecir el corto plazo. Sin embargo, este tipo de modelos se han demostrado ineficientes para la determinación de fenómenos a largo plazo, por lo que quizá sus empleos alternativos deban ceñirse más al campo de la coyuntura económica que al medio y largo plazo.

Cabe destacar algunas conclusiones derivadas del artículo de Nelson (1993): "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model", centrado en que, efectivamente, un proceso ARCH puede dar lugar a una buena estimación de la varianza condicional en tanto en cuanto se pueda aproximar a un proceso de difusión; es decir, en tanto se pueda equiparar a un sistema continuo acotable en intervalos de longitud " $h$ " de forma que, a medida que el valor de " $h$ " tiende a cero, la diferencia entre un valor estimado con cierta carencia de información (infra-especificado) y el valor real de la varianza condicional es cero. Este hecho es fácilmente conseguible cuando se trabaja con datos de muy alta frecuencia, tales como los diarios de los índices, y partiendo de ciertos supuestos que hacen que los procesos ARCH funcionen como filtros de los estimadores de procesos de difusión.

La contraparte de este hallazgo es que, tal y como se ha comentado, hacer "lo bueno con lo malo" está condicionado a procesos con multitud de observaciones donde se pueda simular una diferencia infinitesimal entre dos observaciones, es decir una longitud " $h$ " muy pequeña. Los resultados serán entonces buenos para el corto plazo y, paradójicamente malos en el medio y largo plazo.

A la luz del desarrollo de lo descrito por Nelson, los modelos aplicados resultan útiles para obtener una buena aproximación de la volatilidad incluso si la especificación no fuera perfecta, por lo que cobra especial sentido el desarrollo y explotación de los resultados de esta volatilidad estimada fuera incluso del marco de la función causal a partir de la cual se obtuvieron dichos resultados.

Tal y como se sugería en el último apartado del análisis técnico aplicado, la modelización ARCH es susceptible de mejorar la calidad de algunos indicadores

cualitativos de uso intensivo en el seguimiento de los mercados financieros, tales como el RSI comentado. Evidentemente, todos los indicadores basados en la determinación de cambios probables en la evolución de la tendencia de una serie financiera admiten una mejora inmediata si los valores marcados como puntos de inflexión se calculan de forma más apropiada; es decir, si la desviación típica efectivamente estimada es una buena aproximación de la realidad.

En el caso del ejemplo que desarrollábamos anteriormente, la aplicación de las decilas al recorrido de la serie de la volatilidad nos permite valorar de una forma muy intuitiva el riesgo en el que incurrimos en cada momento en función de los valores registrados por la volatilidad del activo que estemos valorando. En función de la evolución de la volatilidad puntual podemos conocer si ésta representa un valor extremo o no en función de los valores registrados en el pasado.

### **Algunas consideraciones finales**

De todo lo dicho se desprende un camino abierto en la modelización no solo financiera sino de los diversos acontecimientos económicos en los que la estimación de las expectativas juega un papel preponderante.

Quizá se pueda plantear si el esfuerzo modelizador que representa un modelo ARCH incorpora una ganancia en resultados suficiente como para salvar su complejo mecanismo de aplicación. Sobre este particular, señalaría las siguientes cuestiones:

- Las diversas aplicaciones consultadas y la realizada en el contexto de esta tesis para el Ibex-35 nos muestran unos resultados de estimación en el terreno del modelo sobre la media que apenas se ven modificadas por la correcta estimación de la varianza que se sigue del empleo de los modelos tipo ARCH. Los parámetros obtenidos para el modelo de medias móviles apenas cambian.
- Además, la falta de desarrollos conducentes a la eficiencia de la varianza calculada generan dudas razonables sobre el valor estimado con el procedimiento descrito.
- Sin embargo, la validez de la volatilidad condicional como explicativa de la prima de riesgo se pone claramente de manifiesto con este tipo de modelos y,

sólo con ellos, se puede obtener el parámetro que nos indica su importancia exacta en el modelo de regresión planteado.

- Por otro lado, saber que la volatilidad estimada es acertada, incluso con un modelo mal especificado, es una garantía de que su uso aislado - al margen del modelo sobre la media - es una buena aproximación del riesgo (por ejemplo, para realizar un indicador cualitativo como el planteado).
- Por otro lado, el uso frecuente de los modelos ARCH en mercados financieros obliga a realizar un conocimiento apropiado de sus limitaciones y de sus capacidades, de cara a mejorar su aplicación.
- Ha quedado plenamente de manifiesto la bondad de estos modelos para describir el corto plazo, por lo que deben circunscribirse a éste y ser completados con otras especificaciones más de tipo estructural que permitan realizar prospectiva a medio y largo plazo.

En definitiva, el análisis de los modelos ARCH es intrínsecamente positivo por cuanto nos da una visión histórica que de otra forma no sería asumible y que es susceptible de mejorar nuestras perspectivas para el futuro a partir del análisis del pasado. Como todas las herramientas en el terreno de la cuantificación de las Ciencias Sociales, parece patente que es mejor disponer de ellas que no, aunque su utilidad está precisamente en valorar sus resultados convenientemente, de forma que sean una guía a completar por la prudencia de quien la utiliza.

## **ANEXO I: Valores estimados de la varianza condicional autorregresiva y calificación**

	Volatilidad estimada	Decila real	Indicador
4/01/9	2,092	90	más del 75%
5/01/9	2,023	80	más del 75%
7/01/9	1,777	80	más del 75%
8/01/9	1,650	70	-
11/01/9	1,755	80	más del 75%
12/01/9	2,025	80	más del 75%
13/01/9	3,147	100	más del 75%
14/01/9	3,049	100	más del 75%
15/01/9	2,619	90	más del 75%
18/01/9	2,345	90	más del 75%
19/01/9	2,172	90	más del 75%
20/01/9	2,167	90	más del 75%
21/01/9	1,915	80	más del 75%
22/01/9	2,307	90	más del 75%
25/01/9	2,249	90	más del 75%
26/01/9	1,966	80	más del 75%
27/01/9	1,779	70	-
28/01/9	1,551	60	-
29/01/9	1,434	60	-
1/02/9	1,257	40	-
2/02/9	1,153	40	-
3/02/9	1,069	30	-
4/02/9	0,957	20	menos 25%
5/02/9	0,961	20	menos 25%
8/02/9	1,190	40	-
9/02/9	1,376	50	-



10/02/9	1,603	70	-
11/02/9	1,466	60	-
12/02/9	1,283	40	-
15/02/9	1,150	30	-
16/02/9	1,076	30	-
17/02/9	1,033	20	menos 25%
18/02/9	0,929	10	menos 25%
19/02/9	1,026	20	menos 25%
22/02/9	0,952	20	menos 25%
23/02/9	0,892	10	menos 25%
24/02/9	0,871	10	menos 25%
25/02/9	0,812	10	menos 25%
26/02/9	0,871	10	menos 25%
1/03/9	1,101	30	-
2/03/9	1,046	20	menos 25%
3/03/9	0,937	10	menos 25%
4/03/9	1,171	40	-
5/03/9	1,093	30	-
8/03/9	1,032	20	menos 25%
9/03/9	1,135	30	-
10/03/9	1,034	20	menos 25%
11/03/9	0,927	10	menos 25%
12/03/9	0,931	10	menos 25%
15/03/9	0,973	20	menos 25%
16/03/9	0,886	10	menos 25%
17/03/9	0,830	10	menos 25%
18/03/9	0,917	10	menos 25%
19/03/9	0,931	10	menos 25%
22/03/9	0,853	10	menos 25%
23/03/9	1,234	40	-
24/03/9	1,211	40	-
25/03/9	1,076	30	-
26/03/9	1,012	20	menos 25%
29/03/9	1,232	40	-
30/03/9	1,134	30	-
31/03/9	1,188	40	-
6/04/9	1,078	30	-
7/04/9	1,057	30	-
8/04/9	1,091	30	-
9/04/9	0,996	20	menos 25%
12/04/9	1,000	20	menos 25%
13/04/9	1,092	30	-
14/04/9	0,998	20	menos 25%
15/04/9	1,057	30	-
16/04/9	1,386	50	-
19/04/9	1,494	60	-

---

20/04/9	1,343	50	-
21/04/9	1,526	60	-
22/04/9	1,384	50	-
23/04/9	1,211	40	-
26/04/9	1,112	30	-
27/04/9	1,306	50	-
28/04/9	1,210	40	-
29/04/9	1,114	30	-
30/04/9	0,991	20	menos 25%
3/05/9	1,129	30	-
4/05/9	1,051	20	menos 25%
5/05/9	0,983	20	menos 25%
6/05/9	0,881	10	menos 25%
7/05/9	0,818	10	menos 25%
10/05/9	0,774	10	menos 25%
11/05/9	0,708	10	menos 25%
12/05/9	0,693	10	menos 25%
13/05/9	0,649	10	menos 25%
14/05/9	0,637	10	menos 25%
17/05/9	0,994	20	menos 25%
18/05/9	1,045	30	-
19/05/9	0,950	20	menos 25%
20/05/9	0,926	10	menos 25%
21/05/9	0,981	20	menos 25%
24/05/9	0,904	10	menos 25%
25/05/9	0,879	10	menos 25%
26/05/9	0,882	10	menos 25%
27/05/9	1,009	30	-
28/05/9	1,043	30	-
31/05/9	0,950	20	menos 25%
1/06/9	0,866	10	menos 25%
2/06/9	0,818	10	menos 25%
3/06/9	0,751	10	menos 25%
4/06/9	0,748	10	menos 25%
7/06/9	0,689	10	menos 25%
8/06/9	0,651	10	menos 25%
9/06/9	0,685	10	menos 25%
10/06/9	0,643	10	menos 25%
11/06/9	0,893	20	menos 25%
14/06/9	0,851	10	menos 25%
15/06/9	0,998	30	-
16/06/9	0,928	20	menos 25%
17/06/9	0,846	10	menos 25%
18/06/9	0,767	10	menos 25%
21/06/9	0,756	10	menos 25%
22/06/9	0,703	10	menos 25%

---

23/06/9	0,645	10	menos 25%
24/06/9	0,697	10	menos 25%
25/06/9	0,732	10	menos 25%
28/06/9	0,698	10	menos 25%
29/06/9	0,646	10	menos 25%
30/06/9	0,607	10	menos 25%
1/07/9	0,908	20	menos 25%
2/07/9	0,881	20	menos 25%
5/07/9	0,801	10	menos 25%
6/07/9	0,785	10	menos 25%
7/07/9	0,861	20	menos 25%
8/07/9	0,962	30	-
9/07/9	0,955	30	-
12/07/9	0,867	20	menos 25%
13/07/9	1,157	50	-
14/07/9	1,263	60	-
15/07/9	1,401	70	-
16/07/9	1,279	60	-
19/07/9	1,283	60	-
20/07/9	1,161	50	-
21/07/9	1,163	50	-
22/07/9	1,242	60	-
23/07/9	1,358	60	-
26/07/9	1,382	60	-
27/07/9	1,431	70	-
28/07/9	1,294	60	-
29/07/9	1,710	80	más del 75%
30/07/9	1,870	80	más del 75%
2/08/9	1,673	80	más del 75%
3/08/9	1,525	70	-
4/08/9	1,347	60	-
5/08/9	1,688	80	más del 75%
6/08/9	1,617	70	-
9/08/9	1,819	80	más del 75%
10/08/9	1,658	70	-
11/08/9	1,794	80	más del 75%
12/08/9	1,617	70	-
13/08/9	1,413	60	-
16/08/9	1,301	60	-
17/08/9	1,496	70	-

18/08/9	1,368	60	-
19/08/9	1,239	50	-
20/08/9	1,092	40	-
23/08/9	1,034	40	-
24/08/9	1,203	50	-
25/08/9	1,116	50	-
26/08/9	1,016	40	-
27/08/9	1,164	50	-
30/08/9	1,079	40	-
31/08/9	1,363	60	-
1/09/9	1,508	70	-
2/09/9	1,362	60	-
3/09/9	1,829	80	más del 75%
6/09/9	1,691	80	más del 75%
7/09/9	1,683	80	más del 75%
8/09/9	1,562	80	más del 75%
9/09/9	1,374	70	-
10/09/9	1,254	60	-
13/09/9	1,324	60	-
14/09/9	1,456	70	-
15/09/9	1,398	70	-
16/09/9	1,400	70	-
17/09/9	1,438	70	-
20/09/9	1,294	60	-
21/09/9	1,350	60	-
22/09/9	1,315	60	-
23/09/9	1,173	50	-
24/09/9	1,062	40	-
27/09/9	1,203	50	-
28/09/9	1,103	50	-
29/09/9	1,173	50	-
30/09/9	1,105	50	-
1/10/9	1,081	40	-
4/10/9	1,247	60	-
5/10/9	1,159	50	-
6/10/9	1,048	40	-
7/10/9	1,152	50	-
8/10/9	1,066	40	-
11/10/9	0,963	30	-
13/10/9	0,875	20	menos 25%
14/10/9	0,911	20	menos 25%
15/10/9	1,182	60	-

18/10/9	1,251	60	-
			más del
19/10/9	1,397	80	75%
20/10/9	1,280	70	-
			más del
21/10/9	1,416	80	75%
			más del
22/10/9	1,572	90	75%
			más del
25/10/9	1,425	80	75%
			más del
26/10/9	1,599	90	75%
			más del
27/10/9	1,456	80	75%
			más del
28/10/9	1,687	100	75%
			más del
29/10/9	1,546	90	75%
			más del
2/11/9	1,389	80	75%
3/11/9	1,252	60	-
4/11/9	1,136	50	-
5/11/9	1,073	40	-
8/11/9	0,993	30	-
9/11/9	0,925	20	menos 25%
10/11/9	1,066	40	-
11/11/9	0,986	30	-
12/11/9	0,927	30	-
15/11/9	1,018	40	-
16/11/9	0,942	30	-
17/11/9	0,866	20	menos 25%
18/11/9	0,786	10	menos 25%
19/11/9	0,747	10	menos 25%
22/11/9	0,708	10	menos 25%
23/11/9	0,900	20	menos 25%
24/11/9	0,845	20	menos 25%
25/11/9	0,770	10	menos 25%
26/11/9	0,784	10	menos 25%
29/11/9	0,857	20	menos 25%
30/11/9	0,942	30	-
1/12/9	0,872	20	menos 25%
2/12/9	0,824	20	menos 25%
3/12/9	0,767	10	menos 25%
7/12/9	0,744	10	menos 25%
9/12/9	0,753	10	menos 25%
10/12/9	0,940	30	-

13/12/9	0,942	40	-
14/12/9	0,858	20	menos 25%
15/12/9	0,995	40	-
16/12/9	1,021	50	-
17/12/9	0,931	30	-
20/12/9	0,892	30	-
21/12/9	0,811	20	menos 25%
22/12/9	0,796	20	menos 25%
23/12/9	0,728	10	menos 25%
27/12/9	0,695	10	menos 25%
28/12/9	0,854	20	menos 25%
29/12/9	0,903	30	-
30/12/9	0,831	20	menos 25%
3/01/0	0,768	20	menos 25%
			más del
4/01/0	1,396	80	75%
			más del
5/01/0	1,878	100	75%
			más del
7/01/0	2,191	100	75%
			más del
10/01/0	1,982	100	75%
			más del
11/01/0	1,716	100	75%
			más del
12/01/0	1,814	100	75%
			más del
13/01/0	1,775	100	75%
			más del
14/01/0	1,571	90	75%
			más del
17/01/0	1,445	90	75%
			más del
18/01/0	1,284	80	75%
			más del
19/01/0	1,344	80	75%
20/01/0	1,206	70	-
21/01/0	1,236	70	-
24/01/0	1,147	60	-
25/01/0	1,205	70	-
			más del
26/01/0	1,400	90	75%
			más del
27/01/0	1,286	80	75%
28/01/0	1,187	70	-

31/01/0	1,407	90	más del 75%
1/02/0	1,552	100	más del 75%
2/02/0	1,404	90	más del 75%
3/02/0	1,304	80	más del 75%
4/02/0	1,256	70	-
7/02/0	1,236	70	-
8/02/0	1,664	100	más del 75%
9/02/0	1,559	100	más del 75%
10/02/0	1,379	80	más del 75%
11/02/0	1,231	70	-
14/02/0	1,213	70	-
15/02/0	1,538	90	más del 75%
16/02/0	1,604	100	más del 75%
17/02/0	1,425	90	más del 75%
18/02/0	1,272	70	-
21/02/0	1,136	60	-
22/02/0	1,316	80	más del 75%
23/02/0	1,606	100	más del 75%
24/02/0	1,482	90	más del 75%
25/02/0	1,811	100	más del 75%
28/02/0	1,662	100	más del 75%
29/02/0	1,773	100	más del 75%
1/03/0	1,580	90	más del 75%
2/03/0	1,662	100	más del 75%
3/03/0	1,495	90	más del 75%
6/03/0	1,307	70	-
7/03/0	1,171	60	-

8/03/0	1,120	50	-
9/03/0	1,004	40	-
10/03/0	0,990	40	-
13/03/0	1,134	60	-
14/03/0	1,388	80	más del 75%
15/03/0	1,282	70	-
16/03/0	1,257	70	-
17/03/0	1,117	50	-
20/03/0	1,026	40	-
21/03/0	1,006	40	-
22/03/0	1,120	50	-
23/03/0	1,167	60	-
24/03/0	1,295	70	-
27/03/0	1,189	60	-
28/03/0	1,169	60	-
29/03/0	1,053	50	-
30/03/0	1,367	80	más del 75%
31/03/0	1,490	90	más del 75%
3/04/0	1,343	70	-
4/04/0	1,380	80	más del 75%
5/04/0	1,575	90	más del 75%
6/04/0	1,864	100	más del 75%
7/04/0	1,697	100	más del 75%
10/04/0	1,523	90	más del 75%
11/04/0	1,760	100	más del 75%
12/04/0	1,704	100	más del 75%
13/04/0	1,503	90	más del 75%
14/04/0	1,347	70	-
17/04/0	1,466	90	más del 75%
18/04/0	1,416	80	más del 75%
19/04/0	1,257	60	-
20/04/0	1,133	50	-
25/04/0	1,058	40	-



26/04/0	1,096	50	-
27/04/0	1,000	40	-
28/04/0	1,347	70	-
2/05/0	1,264	70	-
3/05/0	1,238	60	-
4/05/0	1,245	60	-
5/05/0	1,174	60	-
8/05/0	1,484	90	más del 75%
9/05/0	1,725	100	más del 75%
10/05/0	1,942	100	más del 75%
11/05/0	2,148	100	más del 75%
12/05/0	1,926	100	más del 75%
15/05/0	1,678	100	más del 75%
16/05/0	1,876	100	más del 75%
17/05/0	1,703	100	más del 75%
18/05/0	1,658	90	más del 75%
19/05/0	1,453	80	más del 75%
22/05/0	1,717	100	más del 75%
23/05/0	1,691	100	más del 75%
24/05/0	1,492	80	más del 75%
25/05/0	1,724	100	más del 75%
26/05/0	1,578	90	más del 75%
29/05/0	1,688	90	más del 75%
30/05/0	1,519	80	más del 75%
31/05/0	1,351	70	-
1/06/0	1,199	50	-
2/06/0	1,127	40	-
5/06/0	1,086	40	-



## ANEXO II: Programas E-views creados para las estimaciones utilizadas en el estudio

### PROGRAMAS SUMINISTRADOS:

#### Modelos tipo ARCH:

- 1.- Modelo I: 'programa para estimar un m-garch(1,1) de regresión con un ma(1) en media
- 2.- Modelo I con distribución t-student
- 3.- Modelo II: M-GARCH(1,1) de regresión con un MA(1) en media y una variable ficticia que discrimina en función del signo de los errores en el período precedente
- 4.- Modelo III.- Modelo II + variable "proxi" sobre rendimientos sin riesgo
- 5.- Modelo IV.- Modelo III + desestacionalización del error
- 6.- Modelo V.- Modelo IV con estructura EGARCH
- 7.- Modelo GJR-ARCH con un proceso de memoria larga (*Component ARCH*)
- 8.- Modelo AP-ARCH
- 9.- Modelo GJR-AP-ARCH

#### Algunos contrastes de ajuste programados:

- 1.- Medidas sobre la capacidad predictiva del modelo
- 2.- Generación de la curva de impacto de las innovaciones
- 3.- Contrastes de sesgo de signo en el modelo estimado

### MODELOS TIPO ARCH

#### Modelo I: 'PROGRAMA PARA ESTIMAR UN M-GARCH(1,1) DE REGRESIÓN CON UN MA(1) EN MEDIA

' Carga del fichero de datos

```
load "c:\views3\ibex35.wf1 "
series y = lnprecio*100
```

' Se establecen dos submuestras: una primera observación para dar los valores iniciales  
' y el resto para la estimación (previamente perdemos un dato por haber generado la  
endógena  
' como diferencia)

```
sample s0 2 2
sample s1 3 2108
smpl s1
```

' Declaración de los vectores de coeficientes que se usan en el logaritmo de máxima  
verosimilitud del ARCH

```
coef(1) mu = .1
coef(1) theta = .1
coef(2) omega = .1
coef(1) alpha = .1
coef(1) beta = .1
```

```
coef(1) gamma=0.2
```

' Se establecen los valores de inicialización de los parámetros a partir del modelo de la media elegido, en nuestro caso con un modelo MA(1)

```
equation eq_temp.ls y c ma(1)
mu(1) = eq_temp.c(1)
theta(1) = eq_temp.c(2)
omega(1) = eq_temp.@se^2
```

' Generación de series adicionales que se emplearán en la maximización

```
smpl s0
series sig2 = omega(1)
series resma = 0
```

' Especificación del logaritmo a maximizar y sus restricciones

```
logl ll1
ll1.append @logl logl
```

' Modelo GARCH(1,1) de regresión y con un MA(1)

```
ll1.append sig2 = omega(1)+alpha(1)*resma(-1)^2 +beta(1)*sig2(-1)
ll1.append resma =y-mu(1) - theta(1)*resma(-1)-gamma(1)*sig2(-1)
```

' Error estandarizado (dividido por la desv típica)

```
ll1.append z = resma/@sqrt(sig2)
```

' Logaritmo de la función normal

```
ll1.append logl = log(@dnorm(z)) - log(sig2)/2
```

' Estimación y muestra de resultados

```
smpl s1
ll1.ml(d)
show ll1.output
```

## **MODELO I CON DISTRIBUCIÓN T-STUDENT**

' Carga del fichero de datos

```
load "c:\reviews3\libex35.wf1 "
series y = lnprecio*100
```

' Se establecen dos submuestras: una primera observación para dar los valores iniciales y el resto para la estimación (previamente perdemos un dato por haber generado la endógena como diferencia)

```
sample s0 2 2
sample s1 3 2108
smpl s1
```

' Declaración de los vectores de coeficientes que se usan en el logaritmo de máxima verosimilitud del ARCH

```
coef(1) mu = .1
coef(1) theta = .1
coef(2) omega = .1
coef(1) alpha = .1
```

```
coef(1) beta = .1
coef(1) tdf = 3
coef(1) gamma=0.2
```

**' para inicializar los parámetros antes de iterar, se emplean los obtenidos con 'la estimación del garch(1,1) gaussiano - suponiendo una distribución normal -**

```
equation eq1
eq1.arch y c ma(1)
show eq1.output
```

```
coef(1) mu = eq1.c(1)
coef(1) theta=eq1.c(2)
coef(1) omega = eq1.c(3)
coef(1) alpha = eq1.c(4)
coef(1) beta = eq1.c(5)
coef(1) gamma=0.2
coef(1) tdf = 3
```

**' Generación de series adicionales que se emplearán en la maximización**

```
smpl s0
series sig2 = omega(1)
series resma = 0
!pi = @acos(-1)
```

**' Especificación del logaritmo a maximizar y sus restricciones**

```
logl ll1
ll1.append @logl logl
```

**' Modelo GARCH(1,1) de regresión y con un MA(1)**

```
ll1.append sig2 = omega(1)+alpha(1)*resma(-1)^2 +beta(1)*sig2(-1)
ll1.append resma =y-mu(1) - theta(1)*resma(-1)-gamma(1)*sig2(-1)
```

**' Error estandarizado (dividido por la desv típica) y grados de libertad - TDF -**

```
ll1.append z = resma^2/sig2/(tdf(1)-2) + 1
ll1.append logl = @gammalog((tdf(1) + 1)/2) - @gammalog(tdf(1)/2) - log(!pi)/2 - log(tdf(1) - 2)/2 -
log(sig2)/2 - (tdf(1)+1)*log(z)/2
```

**' Estimación y muestra de resultados**

```
smpl s1
ll1.ml(d)
show ll1.output
```

**Modelo III: M-GARCH(1,1) de regresión con un MA(1) en media y una variable ficticia que discrimina en función del signo de los errores en el periodo precedente**

**' Carga de datos**

```
load "c:\reviews3\libex35.wf1 "
smpl @all
series y = lnprecio*100
series i=0
```

**' Se establecen dos submuestras: una primera observación para dar los valores iniciales**

**' y el resto para la estimación**

```
sample s0 2 2
sample s1 3 2108
smpl s1
```

**' Declaración de los vectores de coeficientes que se usan en la función máximoverosímil ARCH**

```
coef(1) mu = .1
coef(1) theta = .1
coef(2) omega = .1
coef(2) alpha
alpha(1) = .1
alpha(2)=0.1
coef(1) beta = .1
coef(1) gamma=0.2
coef(1) chi=0
```

**' Valores de inicialización de los parámetros a partir del modelo de la media MA(1)**

```
equation eq_temp.ls y c ma(1)
mu(1) = eq_temp.c(1)
theta(1) = eq_temp.c(2)
omega(1) = eq_temp.@se^2
```

**' Valores muestrales de algunas series necesarias para el desarrollo del logaritmo**

```
smpl s0
series sig2 = omega(1)
series resma = 0
```

**' Declaración de la función verosímil**

```
logl ll1
ll1.append @logl logl
```

**' Modelo GARCH(1,1) de regresión y con un MA(1) y ficticia sobre el residuo del período previo**

```
ll1.append sig2 = omega(1)+alpha(1)*resma(-1)^2 +beta(1)*sig2(-1)+alpha(2)*resma(-1)^2*(RESMA(-1)<0)
```

```
ll1.append resma =y-mu(1) - theta(1)*resma(-1)-gamma(1)*sig2(-1)
```

**' Error estandarizado (dividido por la desv típica)**

```
ll1.append z = resma/@sqrt(sig2)
```

**' Logaritmo de la función normal**

```
ll1.append logl = log(@dnorm(z)) - log(sig2)/2
```

**' Estimación y muestra de resultados**

```
smpl s1
ll1.ml(d)
show ll1.output
```

### **Modelo III.- Modelo II + variable "proxi" sobre rendimientos sin riesgo**

#### **' Carga de datos**

```
load "c:\reviews3\ibex35.wf1 "
smpl @all
series y = lnprecio*100
series j=0
```

#### **' Se establecen dos submuestras: una primera observación para dar los valores iniciales ' y el resto para la estimación**

```
sample s0 2 2
sample s1 3 2108
smpl s1
```

#### **' Declaración de los vectores de coeficientes que se usan en la función máximoverosimil del ARCH**

```
coef(1) mu = .1
coef(1) theta = .1
coef(2) omega = .1
coef(2) alpha
alpha(1) = .1
alpha(2)=0.1
coef(1) beta = .1
coef(1) gamma=0.2
coef(1) chi=0
```

#### **' Valores de inicialización de los parámetros a partir de los del modelo MA(1)**

```
equation eq_temp.ls y c ma(1)
mu(1) = eq_temp.c(1)
theta(1) = eq_temp.c(2)
omega(1) = eq_temp.@se^2
```

#### **' Valores muestrales para algunas series que se utilizan en el logaritmo de verosimilitud**

```
smpl s0
series sig2 = omega(1)
series resma = 0
```

#### **' Declaración del logaritmo máximoverosimil**

```
logl ll1
ll1.append @logl logl
```

#### **' Modelo GARCH(1,1) de regresión con un MA(1), una variable discriminante del signo en el período previo y una variable explicativa adicional**

```
ll1.append sig2 = omega(1)+alpha(1)*resma(-1)^2 +beta(1)*sig2(-1)+chi(1)*lnmibor3m+alpha(2)*resma(-1)^2*(RESMA(-1)<0)
ll1.append resma =y-mu(1) - theta(1)*resma(-1)-gamma(1)*sig2(-1)
```

#### **' Error estandarizado (dividido por la desv típica)**

```
ll1.append z = resma/@sqrt(sig2)
```

#### **' Logaritmo de la función normal**

```
ll1.append logl = log(@dnorm(z)) - log(sig2)/2
```

**' Estimación y muestra de resultados**

```
smpl s1
ll1.ml(d)
show ll1.output
```

**Modelo IV.- Modelo III + desestacionalización del error**

**' Carga de datos**

```
load "c:\views3\libex35.wf1 "
smpl @all
series resma = i
series y = lnprecio*100

series i=0
```

**' Se establecen dos submuestras: una primera observación para dar los valores iniciales  
' y el resto para la estimación**

```
sample s0 2 2
sample s1 3 2108
smpl s1
```

**' Declaración de los vectores de coeficientes que se usan en la función máximoverosímil  
ARCH**

```
coef(1) mu = .1
coef(1) theta = .1
coef(2) omega = .1
coef(2) alpha
alpha(1) = .1
alpha(2)=0.1
coef(1) beta = .1
coef(1) gamma=0.2
coef(1) chi=0
```

**' Valores de inicialización de los parámetros a partir de un MA(1)**

```
equation eq_temp.ls y c ma(1)
mu(1) = eq_temp.c(1)
theta(1) = eq_temp.c(2)
omega(1) = eq_temp.@se^2
genr res=resid
```

**' Valores de inicialización de los parámetros de desestacionalización con una regresión  
ordinaria**

```
equation eq_temp2.ls res c enero diciembre
coef(2) lambda
lambda(1)=eq_temp2.c(1)
lambda(2)=eq_temp2.c(2)
```

**' Generación de algunas series que serán necesarias en el logaritmo de verosimilitud**

```
smpl s0
series sig2 = omega(1)
series resma = 0
```

**' Declaración del logaritmo a maximizar**



```
logl ll1
ll1.append @logl logl
```

**' Modelo GARCH(1,1) de regresión, variable de discriminación del signo de error previo, varianza condicional desplazada explicativa de la media y con un MA(1)**

```
ll1.append sig2 = omega(1)+alpha(1)*resma(-1)^2 +beta(1)*sig2(-1)+chi(1)*lnmibor3m+alpha(2)*resma(-1)^2*(resma(-1)<0)
```

```
ll1.append res =y-mu(1) - theta(1)*resma(-1)-gamma(1)*sig2(-1)
```

```
ll1.append resma =res-lambda(2)*enero-lambda(3)*diciembre
```

**' Error estandarizado (dividido por la desv típica)**

```
ll1.append z = resma/@sqrt(sig2)
```

**' Logaritmo de la función normal**

```
ll1.append logl = log(@dnorm(z)) - log(sig2)/2
```

**' Estimación y muestra de resultados**

```
smpl s1
ll1.ml(d)
show ll1.output
```

### **Modelo V: modelo IV con estructura E-GARCH**

**' Lectura del Workfile**

```
load "c:\reviews3\libex35.wf1 "
smpl @all
```

**' Determinación de la endógena y de una ficticia que se usará posteriormente**

```
series y = lnprecio*100
series i=0
```

**' Se establecen dos submuestras: una primera observación para dar los valores iniciales y el resto para la estimación**

```
sample s0 2 2
sample s1 3 2108
```

```
smpl s1
```

**' Declaración de los vectores de coeficientes que se usan en el ARCH likelihood**

```
coef(1) mu = .1
coef(1) theta = .1
coef(2) omega = .1
coef(2) alpha
alpha(1) = .1
alpha(2)=0.1
coef(1) beta = .1
coef(1) gamma=0.0
coef(1) chi=0
```

**' Valores iniciales a los parámetros sobre regresiones parciales**

```
equation eq_temp.ls y c ma(1)
mu(1) = eq_temp.c(1)
```

```
theta(1) = eq_temp.c(2)
omega(1) = eq_temp.@se^2
genr res=resid
equation eq_temp2.ls res c enero diciembre
coef(4) lambda
lambda(1)=eq_temp2.c(1)
lambda(2)=eq_temp2.c(2)
lambda(3)=eq_temp2.c(3)
```

**' Declaración de series del logaritmo**

```
smpl s0
series sig2 = omega(1)
series resma = 0
```

**' Declaración del logaritmo de máxima verosimilitud**

```
logl ll1
ll1.append @logl logl
```

**' Modelo GARCH(1,1) de regresión y con un MA(1)**

```
ll1.append log(sig2) = log(omega(1))+alpha(1)*resma(-1)^2/@sqrt(sig2(-1))+beta(1)*log(sig2(-1))+chi(1)*lnmibor3m+alpha(2)*(resma(-1)^2/@sqrt(sig2(-1)))*(resma(-1)<0)

ll1.append res = y-mu(1) - theta(1)*resma(-1)-gamma(1)*sig2(-1)

ll1.append resma = res-lambda(2)*enero-lambda(3)*diciembre
```

**' Error estandarizado (dividido por la desv típica)**

```
ll1.append z = resma/@sqrt(sig2)
```

**' Logaritmo de la función normal**

```
ll1.append logl = log(@dnorm(z)) - log(sig2)/2
```

**' Estimación y muestra de resultados**

```
smpl s1
ll1.ml(m=500)
show ll1.output
```

**Modelo GJR-ARCH con un proceso de memoria larga (Component ARCH)**

'Al modelo anterior, que serviría como valores de inicialización de los parámetros, se le añadiría el siguiente desarrollo:

**' Declaración del logaritmo de verosimilitud para el modelo ARCH de memoria larga**

```
logl ll2
ll2.append @logl logl
```

**' Modelo GARCH(1,1) de regresión y con un MA(1). Se incluyen las explicativas anteriores más el proceso diferencial de corto y largo plazo.**

```
ll2.append sig2 = (1-alpha(1)-beta(1))*omega(1)+(phi(1)+alpha(1))*resma(-1)^2 -
(alpha(1)*rho(1)+(alpha(1)+beta(1))*phi(1))*resma(-2)^2+(beta(1)-phi(1))*sig2(-1)+chi(1)*lnmibor3m*100+alpha(2)*resma(-1)^2*(resma(-1)<0)-(beta(1)*rho(1)-
(alpha(1)+beta(1))*phi(1))*sig2(-2)
```

```
ll2.append res =y-mu(1) - theta(1)*resma(-1)-gamma(1)*sig2(-1)
```

```
ll2.append resma =res-lambda(2)*enero-lambda(3)*diciembre
```

**' Error estandarizado (dividido por la desv típica)**

```
ll2.append z = resma/@sqrt(sig2)
```

**' Logaritmo de la función normal**

```
ll2.append logl = log(@dnorm(z)) - log(sig2)/2
```

**' Estimación y muestra de resultados**

```
smpl s1
```

```
ll2.ml(m=500,d,c=0.005)
```

```
show ll2.output
```

## **Modelo AP-ARCH**

**' Carga de datos**

```
load "c:\reviews3\libex35.wf1 "
```

```
series y = lnprecio
```

**' Se establecen dos submuestras: una primera observación para dar los valores iniciales  
' y el resto para la estimación**

```
sample s0 2 2
```

```
sample s1 3 2108
```

```
smpl s1
```

**' Declaración de los vectores de coeficientes que se usan en la función de verosimilitud  
ARCH**

```
coef(1) mu = .1
```

```
coef(1) theta = .1
```

```
coef(1) omega = .1
```

```
coef(1) alpha = .1
```

```
coef(1) gamma=.1
```

```
coef(1) beta = .1
```

```
coef(1) delta=1
```

**' Valores iniciales recogidos del modelo sobre la GARCH(1,1)**

```
equation eq_temp.arch y c ma(1)
```

```
mu(1) = eq_temp.c(1)
```

```
theta(1) = eq_temp.c(2)
```

```
omega(1) = eq_temp.@se^2
```

```
alpha(1) = eq_temp.c(4)
```

```
beta(1) = eq_temp.c(5)
```

**' Creación de algunas series necesarias para el logaritmo**

```
smpl s0
```

```
series sig = @sqrt(omega(1))^delta(1)
```

```
series resma = 0
```

**' Declaración del logaritmo de verosimilitud**

```
logl ll1
```

```
ll1.append @logl logl
```

**' Residuo del modelo MA(1) debido al término constante**

```
ll1.append res = y-mu(1)
```

**' Residuo del MA(1) completo (el de la constante y el del término ma)**

```
ll1.append resma = res - theta(1)*resma(-1)
```

**' Modelo A-PARCH(1,1)**

```
ll1.append sig = (omega(1)+alpha(1)*(@abs(resma(-1)) -gamma(1)*resma(-1))^delta(1)
+beta(1)*(sig(-1))^delta(1))^(1/delta(1))
```

**' Error estandarizado (dividido por la desviación típica muestral)**

```
ll1.append z = resma/(sig)
```

**' Logaritmo de la función normal**

```
ll1.append logl = log(@dnorm(z)) - log(sig^2)/2
```

**' Estimación y resultados**

```
smpl s1
ll1.ml(d)
show ll1.output
```

### **Modelo GJR-AP-ARCH**

**' Carga de datos**

```
load "c:\reviews3\libex35.wf1 "
smpl @all
series y = lnprecio
series i=0
series res=0
```

**' Se establecen dos submuestras: una primera observación para dar los valores iniciales  
' y el resto para la estimación**

```
sample s0 2 2
sample s1 3 2108
smpl s1
```

**' Declaración de los vectores de coeficientes que se usan en el ARCH likelihood**

```
coef(1) mu = .1
coef(1) theta = .1
coef(2) omega = .1
coef(2) alpha
alpha(1) = .1
alpha(2)=0.1
coef(1) beta = .1
coef(1) gamma=0.2
coef(1) chi=0
coef(1) delta=1
coef(1) parfic=1
```

**' Valores iniciales recogidos del modelo sobre la GARCH(1,1)**

```
equation eq_temp.arch y c ma(1)
mu(1) = eq_temp.c(1)
theta(1) = eq_temp.c(2)
omega(1) = eq_temp.@se^2
alpha(1) = eq_temp.c(4)
beta(1) = eq_temp.c(5)
```

**' Valores iniciales de los coeficientes de destacionalización del error MCO**

```
equation eq_temp2.ls res c enero diciembre
coef(3) lambda
lambda(1)=eq_temp2.c(1)
lambda(2)=eq_temp2.c(2)
lambda(3)=eq_temp2.c(3)
```

**' Valores premuestrales para iniciar el logaritmo de verosimilitud**

```
smpl s0
'series sig2 = omega(1)
series sig = @sqrt(omega(1))
series resma = 0
```

**' MODELO ARCH**

```
logl ll1
ll1.append @logl logl
```

**' Modelo  
' Especificación del error**

```
ll1.append res = y-mu(1) - theta(1)*resma(-1)-gamma(1)*sig(-1)^delta(1)
```

**' Desestacionalización del error**

```
ll1.append resma = res-lambda(2)*enero-lambda(3)*diciembre
```

**' Especificación de la ecuación de la desviación típica elevada a delta**

```
ll1.append sig =(omega(1)+alpha(1)*@abs(resma(-1))+alpha(2)*@abs(resma)*(resma(-1)<0)+beta(1)*sig(-1)^delta(1)+chi(1)*lnmibor3m(-1))^(1/delta(1))
```

**' Error estandarizado (dividido por la desv típica)**

```
ll1.append z = resma/sig
```

**' Logaritmo de la función normal**

```
ll1.append logl = log(@dnorm(z)) - log(sig^2)/2
```

**' Estimación y resultados**

```
smpl s1
ll1.ml(500, d)
show ll1.output
```

## ALGUNAS MEDIDAS DE CONTRASTACIÓN EMPLEADAS

### MEDIDAS SOBRE LA CAPACIDAD PREDICTIVA DEL MODELO

```
genr difer=resma^2 - sig2
scalar mse=1/2108*@sumsq(difer)
scalar mae=1/2108*@sum(@abs(difer))
genr diferp=(resma^2 - sig2)/sig2
scalar pmae=1/2108*@sum(@abs(diferp))*100
```

```
genr difer=log(resma^2) - log(sig2)
scalar lmse=1/2108*@sumsq(difer)
scalar lmae=1/2108*@sum(@abs(difer))
```

### GENERACIÓN DE LA CURVA DE IMPACTO DE LAS INNOVACIONES

'Estimación de un E-GARCH y generación de las series de varianzas y residuos

```
smpl 1500 1876
equation eq1
eq1.ARCH(E) lnprecio C
eq1.makegarch garch01
eq1.makeresid resid01
```

```
scalar med = @median(garch01)
```

' GARCH01 es el nombre de la varianza condicional generada

'Generación de una serie - z - que servirá para escribir el eje x de la gráfica de la curva de impacto de las noticias

```
smpl 1 100
series z = -10 + @trend(1)*20/100
```

' Generación de la varianza condicional resultante de los valores estimados

```
series log(sig2E) = eq1.c(2) + eq1.c(5)*log(med) + eq1.c(3)*abs(z) + eq1.c(4)*z
```

'Estimación de un garch y generación de las series de varianzas y residuos

```
smpl 1500 1876
equation eq2
eq2.ARCH lnprecio C
eq2.makegarch garch02
eq2.makeresid resid02
```

```
scalar med2 = @median(garch02)
```

```
smpl 1 100
```

```
series sig2G = eq2.c(2) + eq2.c(4)*(med2)+ eq2.c(3)*z^2
```

'Estimación de un GJR-ARCH (solo con efecto asimétrico) y generación de las series de varianzas y residuos

```
smpl 1500 1876
equation eq3
eq3.ARCH(T) lnPRECIO C
eq3.makegarch garch03
```

```
eq3.makesresid resid03
```

```
scalar med3 = @mean(garch03)
```

```
smpl 1 100
```

```
series sig2T = eq3.c(2) + EQ3.C(3)*Z^2 + eq3.c(4)*Z^2*(z<0) + eq3.c(5)*med3
```

' Gráfica de la curva de impacto de las innovaciones para el modelo E-GARCH, GARCH y GJR-ARCH (sólo término asimétrico)

```
xy Z SIG2E SIG2G SIG2T
```

### **CONTRASTES DE SESGO DE SIGNO EN EL MODELO ESTIMADO**

' debe aplicarse una vez ya se cuenta con la serie del residuo en el modelo ARCH a la que hemos llamado resma

' Generación de una variable ficticia que valga 1 si el residuo precedente fue negativo y cero en el resto de los casos

```
smpl @all
series i=0
series i=1*(resma(-1)<0)
```

' Determinación del sesgo en función de la significatividad del parámetro c(2)

```
equation sesgotest.ls resma/@sqrt(sig2) c i
show sesgotest.output
```

' Determinación del sesgo negativo en función de la significatividad del parámetro c(2)

```
equation sesgonega.ls resma/@sqrt(sig2) c i*resma(-1)/@sqrt(sig2(-1))
show sesgonega.output
```

' Determinación del sesgo positivo en función de la significatividad del parámetro c(2)

```
equation sesgoposi.ls resma/@sqrt(sig2) c (1-i)*resma(-1)/@sqrt(sig2(-1))
show sesgoposi.output
```

' Determinación del sesgo conjunto en función de la significatividad de todos los parámetros - la F-snedecor - .

```
equation sesgoconjun.ls resma/@sqrt(sig2) c i i*resma(-1)/@sqrt(sig2(-1)) (1-i)*resma(-1)/@sqrt(sig2(-1))
show sesgoconjun.output
```

## Referencias Bibliográficas

- AKAIKE, H. (1974): *A new look at the Statistical Model Identification*. IRRR Transactions on Automatic Control, AC-19 716-723.
- BABA, Y, ENGLE, R., KRAFT, D y KRONER, K(1987): *Multivariate Simultaneous Generalized ARCH*. Manuscrito, Universidad de California, San Diego, CA.
- BAILLIE, R. y BOLLERSLEV, T (1992): *Prediction in Dynamic Models with Time-Dependent Conditional Variances*. Journal of Econometrics, 52. Pgs: 91-113.
- BAILLIE, R.T. y BOLLERSLEV, T (1992): *Prediction in Dynamic Models with Time-Dependent Conditional Variances*. Journal of Econometrics, 58, Pgs: 565-585
- BAILLIE, R. y BOLLERSLEV, T. (1989): *The message in Daily Exchange Rates: A conditional Variance Tale*. Journal of Business and Economics Statistics 7. Pgs: 297-305.
- BELSLEY, D. (1979) "On the efficient Computation of Non-linear Full-Information Maximum Likelihood Estimator Paper presentado en en Hallazgos Económicos de la Sociedad Económica, Atenas.
- BERA, A.K.; HIGGINS, M.L. y LEE, S. (1992): *Interaction between Autocorrelation and conditional Heterokedasticity: a Random Coefficient Approach* Journal of Business and Economic Statistics, 10, Pgs.:133-142.
- BERNDT, E. HALL B., HALL R. y HAUSMAN, JA (1974): *Estimation Inference in nonlinear Structural Models* Annals of the Economic and Social Measurement, 4. Pgs: 653-665.
- BOLLERSLEV, T. ENGLE, R. y NELSON, D (1994): *ARCH Models.* Robert Engle y D. McFadden editores. Handbook of Econometrics, Vol IV. Elsevier, Amsterdam.
- BOLLERSLEV, T, CHOU, R y KRONNER, K (1992): *Arch Modelling in Finance: A review of the Theory and Empirical Evidence*. Journal of Econometrics, 52. Pgs: 5-59.



BOLLERSLEV, T. y WOOLRIDGE, J (1992): *Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances*. *Econometric Reviews*, 11.2. Pgs: 143-172.

BOLLERSLEV, T. (1986): *Generalized Autorregresive Conditional Heterocedasticity*, *Journal of Econometrics*, 51. Pgs: 307-327.

BOX PIERCE (1970): *Distribution of Residual Autocorrelations in Autorregresive and Moving Averages Time Series Models*. *Journal of the American Statistical Association*, 65 1509-26.

BREUSCH, T.S. (1978): *Testing for Autocorrelation in Dynamic Linear Models*. *Australian Economics Papers*, 17, pgs. 334-355.

BRUNETI, C. y GILBERT, C. (1999): *Bivariate FIGARCH and Fractional Cointegration*. FEWEC, Vrije Universiteit Amsterdam, GAMW London and CEPR, Documento de trabajo.

CAMPBELL, A. (1993): *Intertemporal Asset Pricing without Consumption Data*. *American Economic Review*, 83. Pgs: 487-512.

CAMPBELL, J. (1987): *Stock Returns and Term Structure*. *Journal of Financial Economics*, 18. Pgs: 373-399.

CHEN, N. ROSS, R. Y ROSS, S. (1986): *Economic Forces and the Stock Market*. *Journal of Business*, 59. Pgs: 383-403.

CHINRONG, A y McFADDEN, D. (1997): *Estimation of some partially specified nonlinear models*. *Journal of Econometrics*, 76. Pgs: 1-37.

DAY, T. y LEWIS, C (1992): *Stock Market Volatility and the Information Content of Stock Index Options*. *Journal of Econometrics*, 52. Pgs: 267-287.

DEMOS, A y SENTANA, E (1998): *Testing for GARCH effects: A one-sided Approach*. *Journal of Econometrics*, 86. Pgs: 97-127.

DICKEY, D y FULLER (1979): *Distribution and Estimators for Autoregressive Time Series with an Unit Root*. Journal of the American Statistical Association, 77. Pgs: 427-431.

DING, Z. y GRANGER, CWJ (1996): *Modeling Volatility Persistence of Speculative Returns: A New Approach*. Journal of Econometrics, 73. Pgs: 185-215.

DING, Z. GRANGER, W.J. y ENGLE, R (1993): *A long Memory Property of Stock Market Returns and a New Model*. Journal of Empirical Finance, 1. Pgs: 83-106.

DOLADO, JENKINSON Y SOSVILLA-RIVERO (1990): *Cointegration and Unit Roots: a survey* Banco de España, Servicio de Estudios. Doc. Trabajo nº 90005.

DROST, F y C. KLAASSEN (1997): *Efficient Estimation in Semiparametric GARCH Models*. Journal of Econometrics, 81. Pgs: 193-221.

DURBIN Y WATSON (1950 y 1951): *Testing for Serial Correlation in Least-Squares Regression*. Biometrika, Vol 37 pgs. 409-428 (1950) y Vol. 38, págs. 159-178 (1951).

ENGLE, R. Y NG, V. (1993): *Measuring and Testing the Impact of News on Volatility*. Journal of Finance, 48. Pgs: 1749-1778.

ENGLE, R. Y GONZALEZ-RIVERA, G. (1991): *Semiparametric ARCH Models*. Journal of Business and Economic Statistics, 9. Pgs: 345-359.

ENGLE, R y GONZALEZ-RIVERA (1990): *Semiparametric ARCH Models*. Manuscript, USCD Department of Economics.

ENGLE, R. NG, V. Y ROTHSCHILD, M. (1990): *Asset Pricing with a FACTOR-ARCH Covariance Structure: Empirical Estimates for Treasury Bills*. Journal of Econometrics, 45. Pgs: 213-237.

ENGEL, C. Y HAMILTON, J (1990): *Long Swings in the dollar: Are they in the Data and Do Markets know it?*. The American Economic Review, September 1990.

ENGLE, R. (1987): *Multivariate ARCH with factor Structures: Cointegration in Variance*. Universidad de California, San Diego, CA.

ENGLE, R.F., LILIEN D.M y ROBINS, (1986): *Estimating the Time Varying Risk Premia in the Term Structure* Econometrica, 55, Pgs.:391-407.

ENGLE, R y BOLLERSLEV, T. (1986): *Modelling the Persistence of Conditional Variance* Econometric Reviews 5, 1-50 y 80-87.

ENGLE, R.F. (1982): *Autorregresive Conditional Heterocedasticity with Estimates of the Variance of the U.K. Inflation* Econométrica, 50. Pgs: 987-1008

FAMA, E y FRENCH, K (1996): *Multifactor Explanations of Asset Pricing Anomalies*. Journal of Finances, 51. Pgs.:55-84

FAMA, y FRENCH, (1993): *Common Risk Factors in the Returns on Stocks and Bonds*. Journal of Finance, 33, Pgs.: 3-56.

FAMA, E. MACBETH, J (1973): *Risk Return and Equilibrium: Empirical Tests* Journal of Financial Economics, 71. Pgs: 607-636.

GALLANT, R. y HSIEH, D. TAUCHEN, G.(1997): *Estimation of Stochastic Volatility Models with Diagnostics*. Journal of Econometrics, 81. Pgs: 159-192

GEWEKE, J. Y PANTULA, S.(1986): *Modelling the Persistence of Conditional Variances: a comment*. Econometric Review, 5. Pgs: 57-61 y 71-74

GIBBSONS, M ROSS S. Y SHANKEN, J (1989): *A Test of the Efficiency of a Given Portfolio*. Econometrica, 57. Pgs: 1121-1152.

GLOSTEN, L . JAGANNATHAN, R. y RUNKLE, D. (1993): *On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks*. Journal of Finance, 48. Num: 5. Pgs: 1779-1801.

GLOSTEN, L. y RUNKLE, D. (1988): *On the Relationship between the Expected Value and the Volatility of the Nominal Excess Returns on Stocks*. Research Dpt. Working Paper, 505. Federal Reserve of Minneapolis.

GODFREY, L.G. (1978): *Testing Against General Autoregressive and Moving Average Models when the Regressors include Lagged Dependent Variables*. Econometrica, 46, pgs. 1293-1302.

GOURIEROUX, C. y MONTFOR, A. (1992): *Qualitative Threshold ARCH Models*. Journal of Econometrics, 52. Pgs: 159-199.

GRANGER, C. y CHOR YIU SIN (1999): *Modelling the Absolute Returns of Different Stock Indices: Exploring the Forecastability of an Alternative Measure of Risk*. University of California San Diego UCSD 9912. Documento de trabajo, diciembre de 1999.

GRANGER, CWJ (1980): *Long Memory relationships and the Agregation of Dynamic Models*. Journal of Econometrics, 14. Pgs: 227-238.

HAMILTON, J. y GANG, L. (1996): *Stock Market Volatility and the Business Cycle*. Journal of Applied Econometrics, 11. Pgs: 573-593.

HAMILTON, J (1994): *Time Series Analysis*. Chap. 22: *Modelling Time Series with Changes in Regime*. Pgs. 677-703. Princeton University Press.

HAMILTON, J (1994): *Contrastes de Especificación en modelos de series temporales markov-cambiantes*. Cuadernos Económicos del ICE, 56. Pgs: 42-70.

HAMILTON, J. Y SUSMEL, R. (1994): *Autorregresive Codnitional Heroskedasticity and Changes in Regime*. Journal of Econometrics, 64. Pgs 307-333.

HAMILTON, JD y SUSMEL, R. (1990): *Autoregressive Conditional Heterocedasticity and Changes in Regime* Journal of Econometrics, 64. Pgs: 307-333.

HAMILTON, J (1989): *A new Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle*. Econometrica, 57.2. Pgs: 357-384.

HAMILTON, J. y SUSMEL, R. (1989): *Autoregressive Conditonal Heteroskedasticity and Changes in Regime*. Journal of Econometrics, 64. Pgs: 307-333.

HAMILTON, J (1988): *Rational Expectations Econometric Analysis of Changes in Regime*. Journal of Economics Dynamics and Control, 12. Pgs: 385-423.

HARVEY, AC (1981): *The Econometric Analysis of Time Series*. Oxford: Phillip Alan.

HAUSMAN, JA (1978): *Specification Test on Econometrics*. Econometrica, 46. Pgs: 1251-1272.

HOYO, J (1993): *Guía para la estimación de modelos ARCH*. Revista Estadística Aplicada, 32. Comentarios. Pgs: 49-57.

KIM, C. (1994): *Dynamic Linear Models with Markov Switching* Journal of Econometrics, 60. Pgs: 1-22.

KRAFT, D. Y ENGLE, R. (1982): *Autoregressive Conditional Heteroskedasticity in Multiple Time Series Models*. USCD, Discussion Paper Núm. 82-23.

LAMOUREUX Y LASTRAPES (1990): *Persistence in Variance, Structural Change and the GARCH Model*. Journal of Business and Economic Statistics, 8. Pgs: 225-234.

LINTNER, J (1965): *The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets* Review of Economics and Statistics, 47. Pgs: 13-37.

L-JUNG-BOX (1978): *On a Measure of Lack of Fit in Time Series Models*. Biometrika, 66 67-72.

MACKINLAY, M. (1987): *On the Multivariate Test of CAPM*. Journal of Financial Economics, 18. Pgs: 341-372.

MARKOWITZ, H. (1959): *Portfolio Selection: Efficient Diversification of Investments*. John Wiley, New York.

Mc CURDY, T. y STENGOS, T (1992): *A Comparison of Risk-premium Forecasts Implied by Parametric versus Nonparametric Conditional Mean Estimators*. Journal of Econometrics, 52. Pgs: 225-244.

MENELAOS, K. (2000): *The Covariance Structure of the S-GARCH and M-GARCH Models*. University of York. Documento de Trabajo Y010 5DD UK.

MILHOJ, A. (1984): *The Moment Structure of ARCH Processes*, Research report 94. Institute of Statistics of Copenhagen, Copenhagen.

NELSON (1993): *Filtering and Forecasting with misspecified ARCH models I: Getting the right variance with the Wrong Model*. Journal of Econometrics, 52. Pgs: 61-90.

NELSON, D.B. (1991): *Conditional Heterocedasticity in asset returns: a New Approach*" Econometrica, 59, Pgs: 347-370

NEUMANN, J. VON (1941): *Distribution of the Ratio of the Mean Square Sucessive Difference to the variance*. Ann. Math Stat. Vol. 12 pgs: 367-395

NEWKEY, W (1985): *Maximum likelihood Specification Test and Conditional Moments Tests*. Econometrica, 53. Pgs: 1047-1069.

PAGAN, A. Y SCHWERT, W. (1990): *Alternative Models for Conditional Stock Volatility*. Journal of Econometrics, 45. Pgs: 267-290.

PANTULA, SG (1986): *Modelling the Persistence of Conditional Variance: A comment*. Econometrica Reviews, 5. Pgs: 71-73.

PANTULA, SG (1984): *Autoregressive Conditionally Heterocedastic Models*. Manuscrito no publicado. Universidad de Carolina del Norte, Dpto. de Estadística.

PEREZ AMARAL, T (1993): *Contrastes m y de la matriz de información dinámica con una aplicación a regresión lineal*. Estadística Española.

PÉREZ, J y C. MURILLO (1997): *Contrastes de Especificación para los modelos de varianza heterocedástica condiconada*" Estudios de Economía Aplicada, 7. Pgs: 101-129.

PERRON, P. (1989): *The Great Crash, the Oil Price Shock and the Unit Root Hypothesis*. Econometrica, 57. Pgs: 1361-1401.

POSKITT, D.S y TREMAYNE, A.R. (1987): *Determining a Portfolio of Linear Time Series Models*. Biometrika, 74, pgs. 125-137.

POTERBA, JM y SUMMERS, LH (1986): *The Persistence of Volatility and Stock Market Fluctuations* American Economic Review, 76. Pgs: 1142-1151.

RICH, R. J. RAYMON y J. BUTLER (1989): *Generalized Instrumental Variables Estimation of Autoregressive Conditional Heteroskedastic Models*. Economics Letters, 35. Pgs: 179-185.

ROSS, S. (1976): *The Arbitrage Theory of Capital Asset Pricing*. Journal of Economic Theory, 13. Pgs: 341-360.

RUÍZ, E. (1993): *Modelos para series temporales heterocedásticas* Cuadernos Económicos ICE, 56, Pgs: 73-108

SCHWARZ, G. (1978): *Estimating the dimension of a Model*. Annals of Statistics, 6 464-464.

SCHWERT, GW. (1989): *Why does Stock Market Volatility Change over Time?* Journal of Finance, 44. Pgs: 1115-1153.

SHARPE, W. (1964): *Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk* Journal of Finance, 19. Pgs: 425-442.

SHIANG, CHIA (1995): *Detecting Parameter Shift in GARCH Models*. Econometric Reviews, 14. Pgs: 241-266.

TAYLOR, S.J. (1986): *Modelling Financial Time Series*, John Wiley, Chichester, U.K..

TAYLOR, S (1986): *Modelling Financial Time Series* Wiley. New York.

TSAY, S.J. (1987): *Conditional Heteroskedastic Time Series Models* Journal of the American Statistical Association, 82. Pgs: 590-604

WEIS, A (1982): *Asymptotic Theory for ARCH Models: Stability, Estimation and Testing*. Discussion Paper 82-36 (Universidad de California, San Diego, CA).

WEST, K. EDISON, H. y CHO, D. (1993): *A Utility Based Comparison of Some Models of Foreign Exchange Volatility*. Journal of International Economics, 35. Pgs: 23-46.

WHITE, H (1987): *Specification Testing in Dynamic Models*. En T. Bewley Advances in Econometric. Vth Congress, Vol 1. Nueva York. Cambridge University Press.

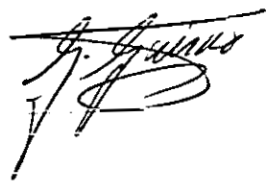
WHITE, H y DOMOWITZ (1984): *Nonlinear Regression with Dependent Observations*. Econometrica, 52. Pgs: 143-161.

ZAKONIAN, J.M.(1994): *Threshold Heteroskedastic Models*. Journal of Economic, Dynamics and Control, 18. Pgs: 931-955.

---



Reunido el Tribunal que suscribe en el día  
de la fecha, acordo calificar la presente Tesis  
Doctoral con la censura de Sobresaliente "con Lode" por unanimidad  
Madrid, 22-1-01



J. B. Peña



A. Pulido

